



HELSINGIN YLIOPISTO
MAATALOUS-METSÄTIETEELLINEN TIEDEKUNTA
METSÄTIETEIDEN LAITOS

METSÄKESKUKSEN HILATIEDON JA LUONNONVARAKESKUKSEN
KOKOJAKAUMAMALLIN TARKKUUS PUUJOUKON LUONNISSA

Matti Pesonen

Pro Gradu- tutkielma

Metsäteknologia ja logistiikka

17.9.2017

Tiedekunta – Fakultet – Faculty Maatalous-Metsätieteellinen tiedekunta		Laitos – Institution – Department Metsätieteiden laitos
Tekijä – Författare – Author Matti Pesonen		
Työn nimi – Arbetets titel – Title Metsäkeskuksen hilatiedon ja Luonnonvarakeskuksen kokojakaumamallin tarkkuus puujoukon luonnissa		
Oppiaine – Läroämne – Subject Metsäteknologia ja logistiikka		
Työn ohjaaja(t) – Arbetets handledare – Supervisor Veli-Pekka Kivinen, yliopistonlehtori Tapio Räsänen, erikoistutkija, Metsäteho oy		Vuosi – År – Year 2017
<p>Tiivistelmä – Abstrakt – Abstract</p> <p>Puuston ennakkotiedon tarkkuuden parantaminen mahdollistaa paremman puunhankinnan suunnittelun, varannon hallinnan sekä kustannusten minimoinnin. Runkolukusarjatasoinen metsävaratieto on muodostumassa puuston keskitunnusten lisäksi yhdeksi keskeiseksi selittäjäksi kuviotasolla puuston koosta, tiheydestä sekä jakaumasta. Metsäkeskuksen tarjoaman hilatiedon sekä Luonnonvarakeskuksen kehittämien Motti-mallien avulla voidaan 16 x 16 metrin hilaruudulle tuottaa teoreettiset runkolukusarjat. Runkolukusarjat tuotetaan ratkaisemalla kahden parametrin Weibull-jakauma hyödyntäen parametrin palautusmenetelmää. Hilan pienipiirteisyyden vuoksi on mahdollista tuottaa otostyyppisiä inventointitapoja tarkempaa ennakkotietoa puuston koosta ja muodostaa monitahoisempia jakaumia, jotka ottavat huomioon kuvion puuston sisäisen vaihtelun.</p> <p>Tarkastelussa hilatiedon ja Motti-mallin pohjalta tuotetun puujoukon tarkkuutta verrattiin päätehakkuiltä kerättyyn hakkuukoneiden tuotantotietoon. Tutkimusala käsitti 150 päätehakkuu- sekä 15 ylispuun poistokuviota, joiden pinta-ala oli vähintään 0,5 hehtaaria. Tuotantotiedon pölkkyistä muodostettiin kokonaiset rungot runkokäyräsovitteen avulla, jonka jälkeen rungot sekä runkokäyräsovitteesta että hilatiedosta simuloitiin katkontasimulaattorissa. Runkolukusarjoja analysoitiin K-S testin ja virheindeksin avulla. Kuvioiden keskitunnuksia pohjapinta-ala, runkoluku, keskipituus ja keskiläpimitta verrattiin ja lopuksi katkottujen puutavaralajien tilavuuksia ja runkojen määrää tutkittiin harhan ja keskineliövirheen neliöjuuren (RMSE) avulla.</p> <p>Tuloksissa K-S testin läpäisi 124 kappaletta 150 päätehakkuukuviosta. Hehtaarikohtaisten puustotunnusten G, N, DG, HG ja V harha 0 %, -14 %, 5 %, 3 % ja 0 % sekä RMSE 18 %, 34 %, 10 %, 5 % ja 17 %. Katkottujen puutavaralajien kokonaistilavuuden harha ja RMSE olivat 1,5 % sekä 16,3 %. Puulajeille mänty, kuusi sekä lehtipuu katkonnan tilavuuksien harhat olivat -72,5 %, 11,5 % ja 2,1 % sekä RMSE 154,5 %, 28,0 %, 50,1 %. Tukki- ja kuitupuiden keskijäreysien harhat olivat 5,9 % ja 3,9 % sekä RMSE 14,0 % ja 10,6 %. Tulosten suurin virhetekijä oli hilatieto. Tarkkuus korreloi parhaiten kuviokoon sekä puuston tasaikäisyyden kanssa.</p>		
<p>Avainsanat – Nyckelord – Keywords</p> <p>Hilatieto, Motti-mallit, puunhankinnan ohjaus, puujoukon ennustaminen, runkolukusarjat, katkonnan simulointi</p>		
<p>Säilytyspaikka – Förvaringsställe – Where deposited</p> <p>Helsingin yliopiston kirjasto – Helda / E-thesis (opinnäytteet) <i>ethesis.helsinki.fi</i></p>		

HELSINGIN YLIOPISTO – HELSINGFORS UNIVERSITET – UNIVERSITY OF HELSINKI

<p>Tiedekunta – Fakultet – Faculty Faculty of Agriculture and Forestry</p>	<p>Laitos – Institution – Department Department of Forest Sciences</p>
<p>Tekijä – Författare – Author Matti Pesonen</p>	
<p>Työn nimi – Arbetets titel – Title The accuracy predicting timber assortments of the grid cell data of Finnish Forest Centre and the diameter distribution model of the Natural Resources Institute Finland</p>	
<p>Oppiaine – Läroämne – Subject Forest technology and logistics</p>	
<p>Työn ohjaaja(t) – Arbetets handledare – Supervisor Veli-Pekka Kivinen, university lecturer Tapio Räsänen, senior reseacher, Metsäteho oy</p>	<p>Vuosi – År – Year 2017</p>
<p>Tiivistelmä – Abstrakt – Abstract</p> <p>Improving the accuracy of forest stand inventory information also improves the efficiency of wood sourcing operations, management of the forest reserve and minimizing the costs. Knowing the diameter-distribution is as important as the traditional stand wise forest attributes. By utilizing the most recent 16 * 16-meter forest grid cell data from the Finnish Forest Centre and the diameter distribution model from the Natural Resources Institute Finland it is possible to simulate the actual tree stems to a specified stand. The diameter distribution model uses the two-parameter Weibull function as well as the parameter recovery method (PRM) in the calculations.</p> <p>The control data was formed from the harvesting data of the same area. The study area consisted of 150 final-felled stands, with the minimum size of 0,5 hectares. The stem data was processed with the taper curve formula to reduce the errors (noise) of the curves and to attach the logs back to stems. Both the stems from the harvesting data and from the simulation were bucked in a bucking simulator. The stand attributes (G, H, DG and HG), diameter distributions as well as the simulated timber assortments were compared through bias and root-mean-squared-error (RMSE). The goodness-of-fit of the diameter distributions was evaluated with the Kolmogorov-Smirnov test and Reynold's Error Index. The visual evaluation of the diameter distributions was also made.</p> <p>In the results 124 of 150 stands passed the Kolmogorov-Smirnov -test. The biases of the stand-level attributes G, N, DG, HG were respectively 0%, -14%, 5% and 0% and RMSE 18%, 34%, 10%, 5% and 17%. The bias and RMSE in total volume of the simulated bucking was 1,5% and 16,3% respectively. Species specified biases in total volume for pine (Pinus sylvestris), spruce (Picea abies) and birch (mixed Betula pendula and B. pubescens) were -72,5%, 11,5% and 2,1% respectively and RMSE 154,5%, 28,0% and 50,1%. The biases of the saw logs and pulpwood logs were 5,9% and 2,1% and the RMSE 14,0% and 10,6% respectively. The major error factor was the grid cell data. The accuracy correlated the best with the size of the stand and the homogeneity of the stems.</p>	
<p>Avainsanat – Nyckelord – Keywords ALS, grid cells, distribution models, wood sourcing, stem prediction, diameter distributions, bucking simulation</p>	
<p>Säilytyspaikka – Förvaringsställe – Where deposited Helsinki University Library – Helda / E-thesis <i>ethesis.helsinki.fi</i></p>	

KIIITOKSET

Tämä työ tehtiin osana Metsätehon projektia 465 *Metsäkonetiedon välitys keskitettyyn tietokantaan ja tietokantasovelluspilotti*, joka kuuluu ”Metsätieto ja sähköiset palvelut” – hankekokonaisuuteen. Projektipäällikkönä toimii Tapio Räsänen. Hankekokonaisuus on osa hallitusohjelman ”Puu liikkeelle ja uusia tuotteita metsästä” – kärkihanketta, jota rahoittaa Maa- ja Metsätalousministeriö. Projektin tavoitteena on luoda perusteet laajamittaiseen metsäkonetiedon kokoamiseen ja hyödyntämiseen puuhuollon sekä metsätalouden tietojärjestelmissä ja käyttösovelluksissa. Uusien menetelmien kehittäminen ja pilotointi hakkuukonetiedon hankkimiseksi, prosessoimiseksi sekä siirtämiseksi tietovarastoihin ja tämän tiedon jakaminen ovat projektin keskeiset tavoitteet.

Tutkielmaa tehtiin pääosin Metsätehon toimitiloissa ja sen toteuttamiseen osallistui tiiviisti Metsätehon tutkijoita. Erikoistutkija Asko Poikelan käsialaa oli runkokäyrän muodostaminen, Tutkija Timo Melkas kehitti puoliautomaattisen kuvionrajausalgoritmin. Erikoistutkija Kirsi Riekkä vastasi molempien käytännön toteutuksesta. Kiitokset heille kaikille työstä, jota en olisi itse pystynyt tekemään. Kiitokset erikoistutkija Tapio Räsäselä laadukkaasta ohjauksesta sekä mielenkiintoisesta gradutyön aiheesta sekä yliopistonlehtori Veli-Pekka Kiviselle niin ikään työn ohjauksesta yliopiston puolesta. Erityiskiitos tutkija Juha-Antti Sorsalle, joka vastasi runkopankkidatasta, kokojakaumamallin käsittelystä sekä katkonnan simuloinnista ja toimitti aineistoja allekirjoittaneelle. Kiitokset kuuluvat myös Luonnonvarakeskuksen tutkija Jouni Siipilehdolle, joka on kehittänyt tutkimuksessa käytetyn kokojakaumamallin sekä tarjosi neuvoa ja apua aina kun sitä tarvittiin.

SISÄLLYS

1 Johdanto.....	1
1.1 Työn tausta	1
1.2 Katkonnan ohjaus ja simulointi	4
1.3 Runkolukusarjamuotoinen metsävaratieto	5
1.4 Optisen kaukokartoitustiedon kerääminen	6
1.5 Laserkeilausaineiston tarkkuus.....	7
1.6 Metsäkeskuksen hilatieto	9
2 Aineisto	11
2.1 Maantieteellinen sijainti	11
2.2 Hakkuukoneaineisto	11
3 Menetelmät	13
3.1 Hakkuukoneaineiston käsittely.....	13
3.2 Hilatiedon hankinta	16
3.3 Kuvioiden rajausta.....	17
3.4 Motti-mallin kuvaus	18
3.5 Puujoukon ennustaminen ja laskenta	20
3.6 Runkolukusarjojen yhteensopivuustestit	22
3.7 Keskitunnusten sekä katkonnan tulosten vertailu	25
3.8 Korrelaatiokerrointen laskenta	28
4 Tulokset	29
4.1 Mallirajoitteet	29
4.2 Kuviorajausten ongelmia.....	30
4.3 Keskitunnusten tarkastelua päätehakkuukuvioilla	31
4.3.1 Pohjapinta-ala	31
4.3.2 Runkoluku	34
4.3.3 Läpimitta	34

4.3.4 Pituus	35
4.3.5 Tilavuus	35
4.3.6 Puulajiosuudet	36
4.3.7 Keskitunnukset	37
4.4 Runkolukusarjojen mallinnus	37
4.4.1 Yhteensopivuustestit	37
4.5 Katkonnan estimoinnin tarkkuus päätehakkuukuvioilla	41
4.6 Puustotunnusten selitettävyys tarkkuuden suhteen	48
5 Tulosten tarkastelua.....	50
5.1 Keskitunnusten virheet	50
5.2 Runkolukusarjat.....	52
5.3 Katkonnan tulokset.....	53
6 Loppusanat	54
Kirjallisuus	56

KÄSITTEET

Pölkky: Yleisnimitys katkotulle puutavaralajille. Voi olla joko mittaan katkottu tukki- tai kuitupuurunko.

Runkolukusarja: Alueellisesti tai muuten rajatun puuston läpimittajakauma rinnankorkeudelta (1,3 m), missä läpimitat ovat pyöristetty luokkiin (esim. 1 cm) ja kunkin luokan läpimittojen lukumäärät esitetään frekvensseinä (Maltamo ja Kangas 1998).

Simulointi: Todellisuutta kuvaavan mallin luominen ja kokeiden tekeminen mallin avulla. Hakkuukoneen katkontaa kuvataan usein simuloinnin avulla, sillä sen monimutkaista katkonnan tulosta on muuten hankala tarkastella (Kivinen ja Uuusitalo 2002).

Solmu: Tietojärjestelmästandardi metsäsuunnitteluun.

STM-tiedosto: StanForD -tiedonsiirtostandardin mukainen tiedosto katkotuista puutavaralajeista (StanForD 2016).

1 JOHDANTO

1.1 Työn tausta

Leimikon ennakkotietojen tarkkuuden merkitys on olennaista puutavaralajien määränhallinnassa. Suomessa puun runko katkotaan valtaosin tavaralajimittaan jo metsässä puun kaatosijainnin läheisyydessä (Uusitalo 1997). Yhdestä rungosta on mahdollista katkoa puuta eri käyttökohteisiin. Rungon lopullinen käyttökohde voi olla sahatavara, vaneri, pylväs tai sellu ja siitä johdettavat lopputuotteet (Kivinen 2004).

Pohjoismaista puunhankintamenetelmää kutsutaan tavaralajimenetelmäksi (Uusitalo 1997), minkä keskeinen vaihe on katkoa puut sopiviin mittoihin kaadon yhteydessä kuljetusten helpottamiseksi. Eri puutavaralajit voidaan suoraan viedä niille varatuille käyttöpaikoille. Tavaralajit voidaan yksinkertaistaen jakaa kahteen luokkaan: sahapuuksi ja kuitupuuksi. Sahapuutavaralajille on olemassa tarkat laatukriteerit sekä pituus- ja läpimittayhdistelmät, jotka asettavat kukin sahalaite erikseen (Piira ym. 2007). Kuitupuulle on määritelty selvästi kevyemmät laatuluokitukset sekä pituusrajoitteet ja käytännössä sen katkontaa eivät ohjaa läpimittarajoitteet. Hakkuukoneissa katkonnan ohjausta ohjaa tietokonejärjestelmä, joka ohjeistaa kuljettajaa katkomaan rungon sopivista kohdista ja maksimoimaan rungosta saatavien tukkitavaralajien määrän (Kivinen ja Uusitalo 2002; Usenius 1980). Eri sahoilla ja vaneritehtailla on omat vaatimukset eri tukkitavaralajeille. Sahat ilmoittavat käyttämänsä pituus- ja läpimittasuhteet puunkorjuusta vastaavalle toimijalle, joka muokkaa matriisimuotoon ja lähettää luodut katkonnanohjaustiedostot hakkuukoneille. Hakkuukone, tai harvesteri, katkoo rungon maksimoiden pölkköjen yhteenlaskettua arvoa. Sahojen pölkyille asettamat laatu- sekä pituus- ja läpimittavaatimukset ovat muodostuneet erilaisten asiakastarpeiden myötä laajaksi kirjoksi (Malinen ym. 2011). Jotta voidaan tietää ennakkoon, mitä puutavaralajeja halutulta metsikkökuviolta syntyy, on siitä oltava riittävän tarkat ennakkotiedot sekä simuloitu katkontatulokset. Suurimmat haasteet katkonnan ohjauksessa ovat riittävän tarkan ennakkotiedon puuttuminen sekä laadun estimointi (Uusitalo 1997). Eri laatutekijöitä, kuten oksaisuutta ja tyvilahoa, ei voida ennustaa kovin hyvin puutasolla ja usein rungon katkontaa joudutaan muuttamaan prosessointihetkellä, mikä muuttaa tavaralajijakaumaa.

Raaka-aineena pohjoismaisella tavaralajimenetelmällä korjatun puun saantoa on hyvin vaikeaa ennustaa etukäteen, johon vaikuttaa pääasiassa kaksi syytä: se on orgaanista hitaasti uudistuvaa raaka-ainetta, jonka kasvuun vaikuttavat ilmastonmuutos, termisen kasvukauden vuosien välinen vaihtelu, luonnonilmiöt ja hyönteistuhot. Toinen syy on pohjoismaisessa hankintamenetelmässä, missä rungot katkotaan tavaralajimittaan metsässä puun kaatopaikalla (Uusitalo 1997). Metsänarviointia on pitkään kehitetty laatimalla alueellisia malleja, joilla puuston kokoa, määrää ja kasvua voidaan ennustaa. Malleilla voidaan arvioida puuston tilavuutta, mutta niillä ei voida määrittää tietyn kuvion puuston tarkkaa runkolukua, tilavuutta, keskiläpimittaa tai puulajisuhdetta metsiköiden sisäisen vaihtelun vuoksi. Tämän vuoksi metsänarviointi on pitkään perustunut otantatyypiseen mittaukseen, jossa mittausten avulla niihin voidaan soveltaa parametrisia tai ei-parametrisia malleja ja siten laskea yhden metsikkökuvion tilavuus. Yksittäisten puiden tunteminen on tärkeää silloin, kun halutaan tietoa puutavaralajijakaumasta (Uusitalo 1997).

Metsävaratiedon tarkkuuteen ovat vaikuttaneet vahvasti taloudelliset intressit. 1980-luvulla käytetty pystymittaus perustui kuvion jokaisen puun läpimitan rekisteröintiin sekä otantapuiden pituuksien laskentaan. Näillä tiedoilla laskettiin koko kuvion puuston tilavuus. Kuvion metsävaratieto oli suhteellisen harhatonta, mutta työlästä tuottaa (Uusitalo 1997, Siipilehto ym. 2016). Seuraavalla vuosikymmenellä yleistyi kuvioittainen arviointi, joka perustui relaskoopilla mitattaviin muuttuvageometrisiin koealoihin. Menetelmä oli nopeampi ja kustannustehokkaampi tuottaa, sillä mittauksiin tarvittiin usein vain yksi maastomittaja. Ilmakuvia hyödyntäen metsikkö jaettiin koostumukseltaan ensin pienempiin, mahdollisimman homogeenisiin, kuvioihin (Poso 2001). Relaskoopilla mitattavat otannat ilmensivät kuvion puuston pohjapinta-alaa, mistä tämän sekä puuston keskipituuden avulla voitiin estimoida puuston tilavuus. Vaikka menetelmä on otettu Metsähallituksessa virallisesti käyttöön vuonna 1907, yleistyi menetelmä myöhemmin myös hakkuiden ennakoarviointiin. Alle 20 vuotta sitten kuvioittainen arviointi oli valtamenetelmä suomalaisessa metsänarvioinnissa (Maltamo ym 2002).

Kuvioittaisen arvioinnin haasteena ovat olleet sen suorittajan subjektiivisuus ja siitä johtuvat virheet. Haaran ja Korhosen (2004) tutkimuksessa todettiin mittaajien välillä harhan olevan -16,2 – 17,0 % sekä virheen olevan 10,6 – 33,9 % keskitilavuuden mittauksessa. Kuvioittaisessa arvioinnissa ei pystytty tarkasti määrittämään kuvion puuston runkolukua, mikä on oleellista runkolukusarjoja muodostettaessa (Siipilehto ym. 2016). Laserkeilauksen edut puustotietojen laskemisessa ovat ylivoimaiset verrattuna muihin nykyisiin maaston inventointimenetelmiin suhteessa käytettyyn työmäärään. Næset (2002,

2004) osoitti tutkimuksissaan, kuinka aluepohjainen menetelmä sopii hyvin suurien metsäalueiden inventointiin.

Hilatieilla viitataan metsävaratietoon, missä kuviotunnukset ovat yleistetty metsässä mitattavaa koealaa vastaavaan alueeseen ja yhdessä ne muodostavat tasoruudukon (Naesset 2001). Hilalla viitataan usein ruudukon yhteen ruutuun, joka sisältää lasketut puustotunnukset. Puustotunnukset ovat luotu kullekin ruudulle yleensä joko k-lähimmän naapurin (k-NN) (Hudak ym. 2008) tai harvalla bayesiläisellä (Junttila ym. 2008) menetelmällä, missä käytetään apuna maastossa erikseen mitattuja referenssikoealoja. Puulajitulkinnassa käytetään apuna laserkeilaustietoa sekä numeerisia ilmakuvia (Packalén ja Maltamo 2007). Referenssikoealoilta mitattujen tarkkojen puustotunnusten perusteella täydennetään puuttuva tieto myös muille hilaruuduille.

Laserkeilaukseen ja ilmakuviin pohjautuvaa aluepohjaista menetelmää on tutkittu ennakkotiedon tarkkuuden määrittämisessä (Naesset 2002, Peuhkurinen ym 2008, Packalén ja Maltamo 2008, Korhonen ym 2008, Holopainen ym 2010, Peuhkurinen ym. 2011, Siipilehto ym 2016). Aluepohjaisella tulkinnalla on päästy parempiin tuloksiin keskitunnuksiin pohjautuviin menetelmiin verrattuna ja sen etuna on kustannustehokkuus sekä leimikon sisällä tapahtuvien muutoksien parempi ennustuskky. Harvapulssista laserkeilausaineistoa ei pystytä hyödyntämään tehokkaasti runkojen laskennassa. Pulssitiheyden olisi oltava vähintään 3 pulssia/m² yksinpuintulkintamenetelmää varten (Peuhkurinen ym. 2011), joten hilakohtainen puusto on ennustettava joko parametrisin (Holopainen ym. 2010, Siipilehto ym. 2016) tai ei-parametrisin (Peuhkurinen ym. 2008, Packalén ja Maltamo 2008) menetelmin. Esimerkiksi Packalénin ja Maltamon (2008) vertailussa ei-parametrisella menetelmällä luodut runkolukusarjat olivat tarkkuudeltaan vain hieman parempia kuin parametrisiin menetelmään perustuneet Weibull-jakaumalla luodut runkolukusarjat.

Suomessa yksityismetsiä inventoi Suomen Metsäkeskus. Metsäkeskuksen metsävaratietojärjestelmällä on hyvät ennakkotiedot yksityismetsistä (Hämäläinen ym. 2017). Vuoden 2016 loppuun mennessä Suomen metsistä oli inventoitu 11,6 miljoonaa hehtaaria eli 82 % yksityismetsien pinta-alasta. Metsäkeskus inventoi vuosittain noin 1,5 miljoonaa hehtaaria metsää, ja koko Suomen inventointikierrrossa kestää siten 10 vuotta. Yksityinen metsänomistaja voi hyödyntää tietoa Metsään.fi – palvelun kautta. Metsänomistaja voi luovuttaa tietoja metsistään halutessaan raakapuun ostajille, jotka voivat tarjota tietojen pohjalta palveluitaan metsänomistajalle. Puun ostajat voivat hyödyntää Metsäkeskuksen tietoa puunhankinnan sekä katkonnan ohjauksessa. Vuoteen

2020 mennessä kaikki Suomen metsät tulisi olla inventoituna. Toisaalta vain 10 % metsänomistajista on rekisteröitynyt palvelun käyttäjiksi. Metsävaratiedon julkistamisen myötä Metsäkeskuksen keräämän tiedon hyödyntämispotentiaali kasvaisi merkittävästi ja loisi perustan metsätietoinfrastruktuurille. Hämäläisen ym. (2017) esittelemässä vuoden 2018 tavoitetilassa Metsäkeskuksen ajantasainen metsävaratieto on keskeisessä asemassa, mitä hyödynnetään metsänhoito- ja puunkorjuutoimenpiteiden ensisijaisena tiedonlähteenä. Kerätyn hakkuukone- ja omavalvontatiedon avulla Metsäkeskuksen metsävaratieto päivitetään ajanmukaiseksi.

Tämän tutkimuksen tarkoituksena on esitellä menetelmä, jossa hyödynnetään Luonnonvarakeskuksen kokojakauma- (Motti-) mallia sekä Metsäkeskuksen hilatietoa, ja jonka avulla voidaan tuottaa estimaatti halutun metsikkökuvion puujoukosta.

1.2 Katkonnan ohjaus ja simulointi

Katkonnan tulos on mielenkiintoinen keino tutkia puujoukkojen vastaavuuksia. Katkonnan ohjaus perustuu joko arvo- ja tavoitematriiseihin tai jakaumamatriisiin, jonka etuna on kevyempi muunneltavuus (Kivinen 2004). Käytännössä nykypäivänä matriiseja ei muunnella leimikkokohtaisesti, jolloin jakauma-aste jää merkittävästi alle sadan prosentin. Tuotantolaitosten asettaman tavoitejakauman sekä toteutuneen jakauman eroa verrataan usein jakauma-asteella, joka ilmoittaa tavoitteen ja toteutuneen pölkkäjakauman välisen suhteen prosenttilukuna. Jakauma-asteen tarkastelu on olennaista tarkastellessa apteerausohjeiden hyvyttä. Ennakkotiedon parantuessa jakauma-aste oletettavasti paranee.

Metsässä tavaralajimittaan katkottu puu voidaan raaka-aineen sijasta olettaa olevan puolivalmiste, jolle asiakas on asettanut ennalta sovittuja mitta- ja laatuvaatimuksia. Esimerkiksi puuta käyttävä sahayritys on antanut tarkat määritelmät sille, mitkä pituus- ja läpimittasuhteet ovat toivottuja vastaanotetuissa pölkyissä ja mitkä laatuluokat pölkyn tulee täyttää (Piira ym. 2007). Pituus- ja läpimittaluokista on luotu koneeseen ohjaustiedosto, matriisi, joka sisältää sallitut pituuden ja läpimitan yhdistelmät (Kivinen ja Uusitalo 2002). Normaalisti tätä katkonnanohjaustiedostoa ei muutella, eikä niiden muuttaminen myöskään onnistu lyhyellä aikavälillä. Dynaaminen katkonnan ohjaus mahdollistaa tuotantolaitosten kysynnän reaaliaikaisen päivittämisen (Hämäläinen ym. 2016). Kasvava tiedonsiirtokapasiteetti sekä korjuuseen kehitettyjen palvelujen ja

ohjelmistojen sisältämä tieto kasvattaa big datan käyttömahdollisuuksia. Edellä mainitun dynaamisen katkonnan esteenä on pitkään ollut liian epätarkka ennakkotieto metsikkökuvion puujoukosta. Tarkemman ennakkotiedon avulla voitaisiin hyödyntää esimerkiksi sumeaa logiikkaa jakaumamatriisien muodostamisessa: simuloitua katkonnan jakaumaa muokataan siten, että se vastaa mahdollisimman tarkasti annettua tavoitejakaumaa ja tämän jälkeen suoritetaan todellinen katkenta hakkuukuviolla (Kivinen ja Uusitalo 2002).

1.3 Runkolukusarjamuotoinen metsävaratieto

Runkolukusarjatasoinen metsävaratieto kuvaa metsikön läpimittajakauman vaihtelua. Koska puun läpimitalla ja pituudella on hyvin vahva korrelaatio, antaa läpimitta myös viitettä puuston kokonaistilavuudesta (Mehtätalo 2007). Läpimittajakaumassa jokainen puu kuvataan kuhunkin läpimittaluokkaan kuuluvana havaintona, jolloin tieto puuston jakaumasta on huomattavasti tarkempaa kuin puuston keskitunnuksiin perustuen. Loetsch ym (1973) esittivät neljä syytä runkolukusarjojen käyttöön: Helppo mitattavuus, luotettavuus, niiden virhelähteet ovat tiedossa sekä näitä voidaan välttää oikeilla välineillä, menetelmillä sekä tarkalla toteutuksella. Lisäksi niistä voidaan johtaa muita tunnuksia, kuten pituus, biomassa sekä tilavuus.

Empiirisiä runkolukusarjoja ei yleensä mitata metsänarvioinnin yhteydessä. Metsänarvioinnissa pitkään tyydyttiin kuvioittaisessa arvioinnissa mitattuihin kuvion keskitunnuksiin, mutta puustonmittausmenetelmät, kuten katkenta- (Uusitalo 1995) tai kasvatussimulaattorit, tarvitsevat lähtötiedokseen runkolukusarjan (Siipilehto 1999). Runkolukusarjat ovat pitkään luotu puuston kuvioittaisista keskitunnuksista saatujen parametrien perusteella. Puuston läpimittajakaumat ovat luotu teoreettisten jakaumien perusteella, kuten Weibull-jakauman, Johnsonin S_B -jakauman (Siipilehto 1999, Siipilehto 2011b) tai prosenttiosuuspisteiden perusteella (Maltamo ym. 2002). Jakauman parametrit ovat laskettu puuston keskitunnuksista, joista tavallisimmin on käytetty puuston runkolukua, keskiläpimittaa tai ikää. Ennustamisen onnistumisen vuoksi on tärkeää mahdollisimman monen puustotunnuksen tunteminen, sillä esimerkiksi runkolukutiedon puuttuminen heikentää ennusteen tarkkuutta (Siipilehto 1999). Puustotunnusten kyky selittää jakauman muotoa on rajallinen, mistä seuraa runkolukusarjojen vahva

keskiarvoistuminen eli runkolukusarjat tapaavat olla samanlaisia kaikissa metsiköissä (Maltamo ym. 2002).

1.4 Optisen kaukokartoitustiedon kerääminen

Optinen kaukokartoitus (LiDAR = Light Detection and Ranging) on tarkimpia metsien inventointimenetelmiä tällä hetkellä (mm. Naeset 2002, Packalén ja Maltamo 2008). Sen etu verrattuna satelliittikuvaukseen on latvuserroksen läpäisevyys ja mahdollisuus kuvata myös alempia kasvuserroksia (Hudak ym. 2008). Kaukokartoituksen toimivuus on todettu paremmaksi havupuuvaltaisissa kuin lehtipuuvaltaisissa metsiköissä, sillä pulssi törmäys puun neulaseen tapahtuu pienemmällä todennäköisyydellä kuin lehteen (Hudak ym. 2008). Se toimii myös tarkasti, vaikka neulasbiomassaa olisi runsaasti. Kaukokartoituksen rajoitteita ovat teknologian sijaan saatavuus ja hinta, jotka ovat rajoittaneet sen operationaalista käyttöä eikä erilaisia hyödyntämismahdollisuuksia ole vielä riittävästi kokeiltu (Hudak ym. 2008). Suomessa harvalla pulssitiheydellä tehtävän lentolaserkeilauksen kustannukset ovat suurissa, koko maan, kattavissa inventoinneissa 15 – 30 €/km² (Holopainen ym. 2013). Laserkeilausaineisto kerätään aktiivisella menetelmällä, jossa käytetty energia, eli laserpulssit, tuotetaan itse. Puustotulkintaa laserkeilausaineistosta suoritetaan pääasiassa kahdella tavalla: Yksinpuintulkintana sekä aluepohjaisella menetelmällä. Yksinpuintulkinta on tarkempaa, mutta myös kalliimpaa. Tässä menetelmässä laserkeilaimen pistetiheys on noin 5-20 pulssia /m², mutta miniminä voidaan pitää kolmea pulssia neliömetrillä (Peuhkurinen ym. 2011). Aluepohjaisessa menetelmässä puustotiedot yleistetään hilaruudulle, jonka kokona on yleisimmin käytetty Suomessa 16 m * 16 m vastaten siten referenssikoealan kokoa. Pistetiheytenä kyseisessä menetelmässä on käytetty 0,5-2 pulssia/m². Optisen kaukokartoituksen perusteella ei voida tehdä puulajitulkintaa, vaan se täytyy tehdä ilma- tai satelliittikuvien perusteella (Hudak ym. 2008).

Kaukokartoitustiedon tuottamiseen tarvitaan aineiston lisäksi paljon tilastolliseen laskentaan perustuvia toimenpiteitä. Tällä hetkellä laserkeilausaineistoa hyödynnetään metsänarvioinnissa ja pulssien muodostamasta pisteaineistosta on erilaisin menetelmin laskettavissa monenlaisia erilaisia puustotietoja (Hudak ym. 2008). Puu- ja puustotunnusten laskemiseen tarvitaan referenssiaineistoa maastosta, jotka toimivat selitettävänä muuttujina pistepilven selittäville muuttujille, kun käytetään ei-parametrisia

menetelmiä. Laserkeilausaineiston tarkkuutta puusto- ja tilavuustunnusten välillä on selitetty aiemmin lähinnä harhan (bias) ja keskineliövirheen neliöjuuren (RMSE) avulla (Vauhkonen ym. 2013).

K-lähimmän naapurin (k-NN) menetelmä on ei-parametrisen regression yksi muoto, joka muodostaa parametrisen regression tavoin vastemuuttujien muodostuksessa selittävien muuttujien arvoja (Hudak ym. 2008). Toinen yleisesti käytetty puustotunnusten laskentamuoto on harva bayesiläinen (Sparse Bayesian) menetelmä (Junttila ym. 2008). Vastemuuttujien muodostusta varten tarvitaan maastosta maastokoealoja, joilta mitataan sekä vastemuuttujien että selittävien muuttujien arvot. Kaukokartoitusaineistosta irrotetaan vain selittävien muuttujien piirteet. Maastokoealojen puuston oletetaan edustavan koko alueen puustoa, joten selittäjien ja vastemuuttujien välisten korrelaatioiden avulla muodostetaan vastemuuttujille arvot koko kaukokartoitusalueelle. Ei-parametrisessa laskennassa virheet ovat aina suurempia kuin regressiolaskennassa ja osasyynä tähän on, että ennustetut arvot ovat laskettu havaintojen joukosta; tässä tapauksessa mitatuista koealoista. Virhettä voidaan laskea keskineliövirheen neliöjuuren (RMSE) avulla.

1.5 Laserkeilausaineiston tarkkuus

Aiemmissa tutkimuksissa on aluepohjaista menetelmää käyttäen päästy hyvään puustotarkkuuteen. Seuraavaksi käsiteltävissä tutkimuksissa on referenssiaineistona käytetty hakkuukoneen tuotantotietoja. Siipilehdon ym. (2016) tutkimuksessa vertailtiin erilaisia inventointimenetelmiä toteuttaa runkolukusarja kuviotasolla. Siinä todettiin hilaruudukkomuotoisen laserkeilaustiedon olevan yksi tarkimmista nykyaikaisista menetelmistä puustotiedon mittauksessa. Tutkimuksessa todettiin myös, että mikäli ajantasaista aluepohjaista laserkeilausaineistoa on saatavilla kohteesta, tuovat maastokoealamittaukset etenkin havumetsiköissä vain vähän lisäarvoa puustotietoon. Siipilehdon ym. (2016) tutkimuksen lähtötietoaineiston muodostivat seitsemän kuviota Hämeenlinnan, Orimattilan sekä Myrskylän kunnissa, jotka olivat mäntyvaltaisia tai koivusekapuumetsiä. Siipilehdon ym. (2016) tutkimuksessa aluepohjaisen tiedon luonnissa käytettiin RandomForest (RF) menetelmää lähimmän naapurin (NN) valinnassa. Lähimpiä naapureita valittiin vain yksi joka hilalle.

Holopainen ym. (2010) tutkivat laserkeilauksella tehtävän inventoinnin epätarkkuutta tilavuuden, puutavaralajien sekä puukauppatulojen arvoihin. Aineisto koostui 31

harvennus- ja uudistushakkuualasta.. Tutkimuksessaan he totesivat, että puuston ennakkotiedon keruulla oli suurin merkitys puujoukon tarkkuuden määrittämiseen, kun virheitä vertailtiin rungon muodostamisen ja katkonnan kanssa. Tutkimuksessa mainitaan myös virheellisen ennakkotiedon merkitys hakkuuajankohdan määrittämiseen, mistä voi tulla tappioita hakkuusaantoon. Tutkimus todisti myös kaukokartoituksella kerätyn metsävaratiedon olevan paljon tarkempaa kuin kuvioittaisen arvioinnin.

Peuhkurisen ym. (2008) tutkimus järjestettiin metsäyhtiö UPM Kymmene Oyj:n omistamissa metsissä ja keilausala oli yhteensä noin 1200 ha. Peuhkurisen ym. (2008) tutkimuksessa verrattiin laserkeilaustiedon, ilmakuvatulkinnan sekä näiden yhdistelmän tarkkuutta puuston tukkitilavuuden määrittämisessä. Tutkimuksessa käytettiin referenssiaineistona hakkuukonetietoa ja tukkipuun saantoprosenttia estimoitii tutkimusaineistosta saadulla ja erillisellä runkopankkiaineistolla. Tutkimuksessa todettiin, että tukkisaannon osuuden määrittäminen ei onnistu yhtä helposti kuin tukkitilavuuden laskeminen, sillä runkojen katkaisukohdat eivät aina sijaitse samassa kohtaa kuin pölkyn minimiläpimitan korkeus. Korhonen ym. (2008) päätyivät samanlaisiin tuloksiin tutkimuksessa, jossa tutkittiin puuston tukkiosan estimoinnin tarkkuutta; RMSE oli 19,2 m³/ha (9,1 %), mutta todellisen saannon RMSE oli 30,6 m³/ha (18,0 %).

Packalén ja Maltamo (2008) esittivät tutkimuksessaan ei-parametrisen ja parametrisen menetelmän eroja runkolukusarjan luonnissa kaukokartoitustiedosta. Ei-parametrisessa menetelmässä referenssikoealojen puustotiedot toimivat suoraan hilatiedon selittäjinä, kun hiloille etsittiin lähimpiä naapureita. Lähimpien naapurien lukumääränä käytettiin tutkimuksessa viittä. Kun näiden menetelmien tarkkuutta verrattiin, huomattiin ei-parametrisella menetelmällä laskettujen puujoukkojen olevan selkeästi vähemmän harhaisia kuin parametrisella menetelmällä laskettujen, vaikka virhe oli molemmilla samansuuruinen. Tutkimuksessa sovellettiin parametrin ennustamismenetelmää (PPM). Sekä Peuhkurisen ym. (2008) että Packalénin ja Maltamon (2008) tutkimuksissa käytetyn keilaimen pulssitiheys oli 0,7 pulssia/m², mikä on hyvin yhdensuuntainen Metsäkeskuksen käyttämän pulssitiheyden kanssa. Hakkuukoneen tuotantotiedot kerättiin yhteensä 14 kuviolta, jotka olivat niin ikään havupuuvaltaisista.

Peuhkurisen ym. (2011) tutkimuksen tarkoituksena oli tutkia aluepohjaisen menetelmän ja yksinpuintulkinnan eroja puuston tilavuuden ennustamisessa Kolin kansallispuiston alueella. Pulssin pistetiheys oli neljä pulssia neliömetrille, jotta samaa aineistoa voitiin käyttää molempiin laskentamenetelmiin. Runkolukusarjan ennustaminen hilatiedosta tehtiin parametrin palautusmenetelmää hyödyntämällä, jotka ennustivat

läpimitta- sekä pituusjakaumat koealoille joko puuston runkoluvun tai puuston pohjapinta-
alan avulla. Mehtätalon ym. (2007) mukaan. Aluepohjaisen menetelmän kohdalla
runkoluvun estimointi oli melko harhatonta, vaikka virhettä esiintyi. Tutkimuksen
tuloksissa todettiin aluepohjaisen menetelmän ja parametrin palautusmenetelmän
yhdistelmän aliarvioivan suurien puiden määrää ja toimivan paremmin kuvioilla, jonka
runkolukusarja oli painottunut pieniin läpimittoihin. Perinteisten harhan ja keskivirheen
lisäksi tutkimuksessa laskettiin Pearsonin korrelaatiokertoimet eri puustotunnusten sekä
ennustustarkkuuksien välille. Korrelaatio oli selkeintä puuston aritmeettisen keskiarvon ja
pituuden sekä kokonaistilavuuden virheen välillä. Myös painotettua Reynolds'n
virheindeksiä käytettiin runkolukusarjojen arvottamiseen.

1.6 Metsäkeskuksen hilatieto

Suomen Metsäkeskus kerää metsävaratietoa pääosin kahdella eri menetelmällä;
maastossa tehtävällä kuvioittaisella arvioinnilla sekä kaukokartoitusperusteisesti.
Metsäkeskuksen kaukokartoituspohjainen metsävaratieto perustuu laserkeilaukseen,
ilmakuvaukseen, referenssikoealojen mittaukseen sekä aineistojen mukaan suoritettavaan
tilastollisiin laskentoihin (Suomen Metsäkeskuksen metsävaratiedon laatuseloste 2016).
Menetelmä otettiin käyttöön vuonna 2010 ja siihen siirryttiin kokonaan seuraavana
vuonna. Tiedon luotettavuuteen vaikuttavat tiedonkeruumenetelmät, tietolähteet sekä
tiedon ajankohta. Kaukokartoituksellisesti mitatut haastavat kohteet, kuten taimikot,
uudistusalat tai epätasaiset metsiköt, mitataan osittain edelleen maastossa. Tietoa voidaan
täydentää myös eri toimijoiden metsäsuunnittelutiedolla.

Metsäkeskuksen kaukokartoitustieto kerätään harvalla pulssitiheydellä.
Yksinpuintulkintaa varten kerättävän laserkeilauksen pulssitiheys tulisi olla vähintään
keskimäärin kolme pulssia neliometriä kohden (Peuhkurinen ym. 2011), jolloin aineiston
tuottaminen on myös kalliimpaa. Metsäkeskuksen tilaaman kaukokartoitustiedon
pulssitiheys on noin yksi pulssi kahta neliometriä kohden (Suomen Metsäkeskuksen
metsävaratiedon laatuseloste 2016), jolloin aluepohjainen menetelmä on luonnollinen
vaihtoehto.

Metsäkeskuksen tuottama metsävaratieto on ensisijaisesti suunnattu yksityisille
metsänomistajille (Suomen metsäkeskuksen metsävaratiedon laatuseloste 2016).
Vuositain Metsäkeskus inventoi noin 1,5 miljoonaa hehtaaria yksityismetsiä Suomessa.

Ala käsittää keskimäärin noin 10 - 12 aluetta, joista kullekin määritellään 600 - 800 referenssikoealaa. Keskimäärin referenssikoealoja sijoitetaan siten yksi 200 hehtaaria kohden. Metsänomistaja voi ladata tietoa metsistään Metsään.fi – palvelun kautta. Myös julkishallinnolliset metsänomistajat, kuten kunnat tai seurakunnat, voivat hyödyntää tietoa, mutta näille alueille ei tehdä referenssikoealamittauksia.

Referenssikoealojen mittauksissa yhden ympyräkoealan läpimitta on kahdeksantoista metriä, harvapuustoisessa metsikössä 25,24 ja taimikoissa koealalle sijoitetaan useita pienempiä koealoja. Alalta mitataan kaikki merkittävät puut ja koealan keskipiste paikannetaan vähintään metrin tarkkuudella (Suomen Metsäkeskuksen metsävaratiedon laatuseloste 2016). Siemen- ja ylispuukuvioilla ei inventoida suurempia ylis- tai siemenpuita, vaan kyseisen puun osuessa koealalle koealaa siirretään (Kaukokartoitusperusteisen metsien inventoinnin... 2014).

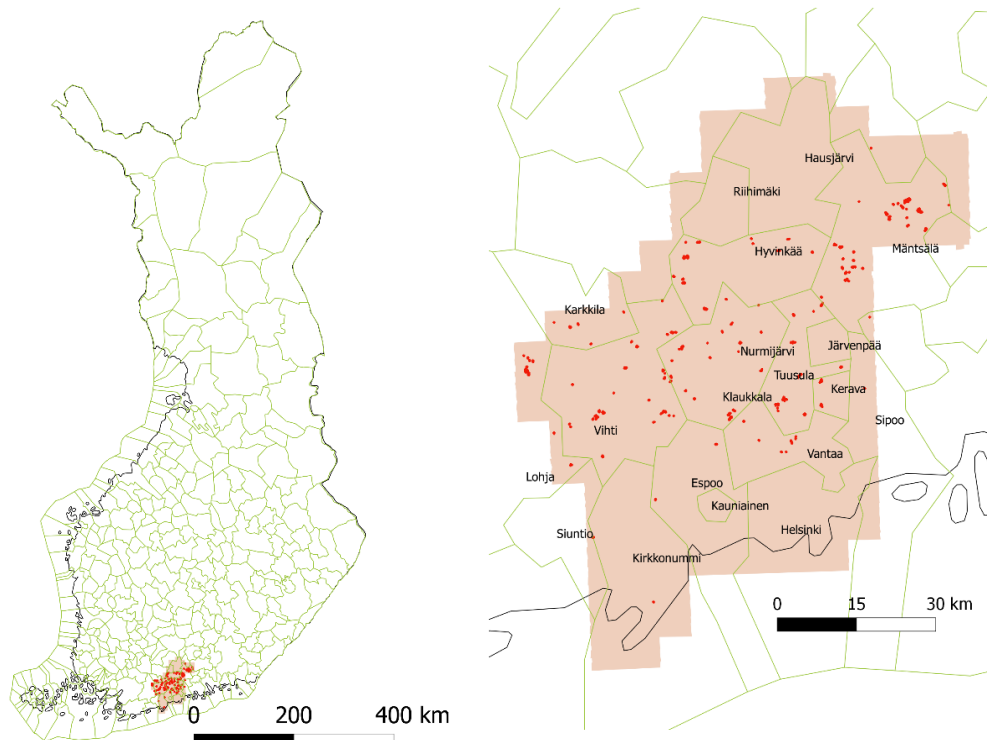
Metsäkeskuksen hilatiedon etuna on sen kattavuus. Koska Metsäkeskus tuottaa laserkeilaustietoa yksityismetsänomistajien käyttöön valtakunnallisesti, on Metsäkeskuksen laserkeilausaineiston hyödyntämisessä paljon mahdollisuuksia, joista runkolukusarjainen metsävaratieto on yksi. Tiedon kattavuus, julkisuus ja helppo saatavuus ovat edellytyksenä menetelmän käyttöönottoon: Kun metsänomistaja antaa suostumuksen, tietoa voidaan hyödyntää myös leimikkosuunnittelussa tai muissa metsänhoidollisissa toimenpiteissä.

Laserkeilauksen hilatietoon liittyy virhetekijöitä. Vaikka laserkeilaustieto toimii erityisen hyvin havupuuvaltaisissa metsiköissä, joissa neulaset eivät estä pulssia kulkeutumasta maahan, ei puulajitulkintaa voida tehdä pelkän laserkeilausaineiston perusteella (Hudak ym. 2008). Tämän vuoksi puulajitulkinta tehdään numeeristen ilmakuviin perusteella. Siitä huolimatta etenkin havupuulajien erottaminen toisistaan tuottaa vaikeuksia. Puulajiosuuksien väärintulkinnassa myös puutavaralajien osuudet vaihtelevat, mistä seuraa virhettä ennustukseen.

2 AINEISTO

2.1 Maantieteellinen sijainti

Tutkimusalue sijoittui Länsi-Uudenmaan alueelle Etelä-Suomeen. Tutkimusalue koostui niistä alueista, joista oli saatavilla sekä Metsäkeskuksen kaukokartoitustietoa sekä hakkuukoneen tuotantotietoa. Alueen puusto oli pääasiassa metsäkuusta (*Picea abies*) (76,2 %), koivua (*Petula bendula* ja *P. pubescens*) (11,9 %) sekä mäntyä (*Pinus sylvestris*) (9,0 %). Muiden lehtipuiden osuus oli 3,0 %. Kuusi oli pääpuulajina kaikkiaan 82 % kuvioista. Puuston yhteenlaskettu kokonaistilavuus oli tutkimuskuvioilla 89 539 m³.



Kuva 1. Tutkimusaineisto sijaitsi Länsi-Uudenmaan alueella, minkä muodosti Metsäkeskuksen laserkeilausaineisto (alue) sekä pääte- ja yleispuuhakkuiden tuotantotiedot (pist). Pohjakartta: Yleiskartta, Hallintorajat ©Maanmittauslaitos 04/2017

2.2 Hakkuukoneaineisto

Hakkuukoneaineiston muodostivat Forest Big Data ja Katkonnan ohjaus -hankkeisiin kerätyt ja sieltä poimitut hakkuiden tuotantotiedot Länsi-Uudenmaan alueelta (Kuva 1). Hakuut olivat suoritettu keilauksen jälkeen vuosien 2015 ja 2016 aikana. Puuston laserkeilauksen ja hakkuiden ajallinen väli oli lyhyt, joten puiden kasvua ei tarvinnut

simuloida. Hakkuukoneaineisto toimi tutkimuksessa havaittuna aineistona. Aineisto sisälsi kaikki projekteihin kerätyt hakkuun tuotantotiedot, ja sieltä poimittiin tätä tutkimusta varten kaikki laserkeilausalueelle osuneet hakkuut, joissa jäävä puusto jäi mahdollisimman pieneksi. Siten aineistoon valikoituivat kaikki päätehakkuuleimikot.

Hakkuukoneen tuotantotiedosta oli saatavilla alkuperäisten runkojen dimensiot. Hakkuukonemittaus on suhteellisen luotettava mittaustapa, mutta paikallista virhettä syntyy dimensioihin mm. oksien ja haaraisuuksien vuoksi. Tämän vuoksi hakkuukoneella puidut rungot ovat muodostettu uudelleen runkokäyräsovitteen avulla, jolloin rungoista saadaan muodostettua mahdollisimman oikeanmallinen. Runkokäyräsovitteen avulla voitiin luoda todellinen puujoukko mahdollisimman tarkasti. Puulaji määriteltiin hakkuukoneenkuljettajan syöttämän tiedon mukaan, ja tässä tutkimuksessa puulajit jaoteltiin kolmeen luokkaan: männyt (1), kuuset (2) sekä lehtipuut (3).

Taulukko 1 Päätehakkuukuvioiden keskitunnukset hakkukonemittausten perusteella (n = 150). DG ja HG tarkoittavat pohjapinta-alalla painotettuja keskiläpimittaa (D) ja pituutta (H), SD tarkoittaa keskihajontaa sekä CV variaatiokerrointa.

	Keskiarvo	Minimi	Maksimi
Kuvion pinta-ala, ha	2.2	0.5	12.2
Hakkuutapa	4.0	4.0	4.0
Tyyppi	1.0	1.0	1.0
Pohjapinta-ala, m ² /ha	25.0	8.4	41.3
Runkoluku, kpl/ha	591.0	224.4	1077.2
DG, cm	28.7	14.9	46.8
HG, m	22.8	16.1	29.5
Puuston tilavuus, m ³ /ha	263.9	63.6	526.8
Rungon keskijäreys, m ³ /ha	0.48	0.12	1.10
SD läpimitta, cm	8.5	4.1	12.2
SD pituus, m	5.3	3.4	8.1
CV läpimitta	0.39	0.29	0.54
CV pituus	0.28	0.19	0.38
Osuus mänty, %	10 %	0 %	76 %
Osuus kuusi, %	72 %	2 %	100 %
Osuus lehtip., %	18 %	0 %	98 %

3 MENETELMÄT

3.1 Hakkuukoneaineiston käsittely

Hakkuukoneaineiston STM-tiedostojen rungot käsiteltiin runkokäyräsovitteella, jotta rungoista voitiin muodostaa selkeämpi ja vähemmän häiriöinformaatiota sisältävä malli (Poikela, henkilökohtainen tiedonanto, kuva 2). Runkokäyräsovitte on tehty Laasasenahon (1982) tutkimuksessa esitetyn yhtälöä mukaillen. Runkokäyräsovitte tehdään C++ ohjelmalla ja lähtöaineistona on hakkuukoneen tuottama STM-tiedosto JSON-formaatissa (Riekkö, henkilökohtainen tiedonanto). Tarvittavat parametrit ovat rungon numerotiedot ja mahdolliset osarungon numerot, puulaji, käyttöosan pituus, ylimmän pölkyn latvaläpimitta, tukkiosan pituus sekä ylimmän tukin latvaläpimitta, joita käytetään runkokäyrän sovitukseen sekä tilavuuslaskentaan. Runkokäyräsovitte ei varsinaisesti tarvittu tätä tutkimusta varten, vaan tässä tutkimuksessa käytettiin runkokäyräsovitteella korjattuja puita, minkä vuoksi sovitteen menetelmä on esitetty lyhyesti.

Sovitukseen käytetään STM-tiedostosta löytyvää puun 12. läpimittaa d_{11} , josta lasketaan puun pituus. Läpimitta d_{11} tarkoittaa kaatoleikkauksen jälkeen kymmenen sentin välein poimittua 12. läpimittaa. Senttimetreissä ilmoitettuna läpimitan mittauskorkeus on 120 cm kaatoleikkauksesta. Kullekin puulajille (mänty, kuusi ja koivu) sovelletaan omaa laskentakaavaa puun korkeuden määrittämiseksi. Kannonkorkeuden laskennassa käytetään apuna puun pituutta h sekä läpimittaa d_{11} ja kullekin puulajille on laskentaan käytössä oma kaava.

Keskimääräinen puun pituus $h(m)$ laskettiin seuraavasti (käytettiin sekä lähtöarvoissa että rajoitteissa):

$$(3.1) \text{ Mänty: } h(m) = \text{Min} [26m ; 1.3 + 0.993 * d_{11}(mm)/10 - 0.00907 * (d_{11}(mm)/10)^2]$$

$$(3.2) \text{ Kuusi: } h(m) = \text{Min} [26.5m ; 1.3 + 0.993 * d_{11}(mm)/10 - 0.00907 * (d_{11}(mm)/10)^2]$$

$$(3.3) \text{ Koivu: } h(m) = \text{Min} [26m ; 1.3 + 0.993 * d_{11}(mm)/10 - 0.00907 * (d_{11}(mm)/10)^2]$$

Kannonkorkeuden $h_k(m)$ lähtöarvot laskettiin seuraavasti:

$$(3.4) \text{ Mänty: } h_k(m) = \text{Max} [0.000952 * h(m) + 0.004456 * d_{11}(mm)/10 ; 0.07m]$$

$$(3.5) \text{ Kuusi: } h_k(m) = \text{Max} [0.0056 * h(m) + 0.005089 * d_{11}(mm)/10 ; 0.07m]$$

$$(3.6) \text{ Koivu: } h_k(m) = \text{Max} [0.004979 * h(m) + 0.004862 * d_{11}(mm)/10 ; 0.07m]$$

Sovituksessa käytettävä kaava on esitetty Laasasenahon (1982, s. 22) tutkimuksessa

$$(3.7) \frac{d_1}{d_{2h}} = b_1x + b_2x^2 + b_3x^3 + b_4x^5 + b_5x^8 + b_6x^{13} + b_7x^{21} + b_8x^{34}$$

, jossa d_{2h} on rungon läpimitta 20 % korkeudella maanpinnasta, d_1 on rungon ensimmäinen läpimitta sekä $x = 1 - \frac{1}{h}$ tai suhteellinen korkeus latvasta. Yhtälön potenssit noudattavat niin kutsuttua Fibonaccin lukujonoa. Runkokäyrän sovitus koostuu pääpiirteissään neljästä osiosta. Päämääränä on löytää STM-tiedoston yksittäistä runkoa vastaavat tiedot puun kokonaiskorkeudesta h , kannon korkeudesta h_k , läpimitasta 20 % osakorkeudella $d_{0,2h}$ sekä runkoa mallintavan polynomi yhtälön kolme ensimmäistä b -parametria. Sovitukseen käytetään Levensberg-Marquardt -algoritmia (Lourakis 2005). Algoritmi toimii pienimmän neliösumman periaatteella.

Ensimmäisessä vaiheessa em. parametreille annetaan karkeat arvaukset, jotka perustuvat Laasasenahon tutkimuksessa (1982, s. 28) annettuihin puulajikohtaisiin b :n arvoihin (3.8).

		b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8
(3.8)	Mänty	2,1288	0,63157	-1,6082	2,4886	-2,4147	2,3619	-1,7539	1,0817
	Kuusi	2,3366	-3,2684	3,6513	-2,2608	0,0	2,1501	-2,7412	1,8876
	Koivu	0,93838	4,1060	-7,8517	7,8993	-7,5018	6,3863	-4,3918	2,1604

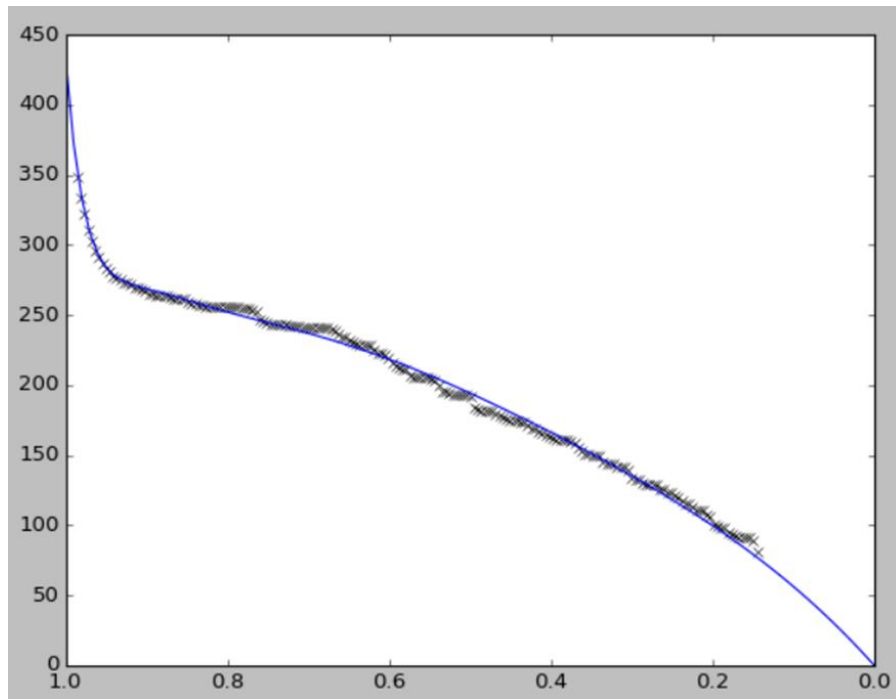
Sovitusta tarkastellaan rajoitteiden avulla, jotka pitää asettaa itse. Laskennoissa minimikannonkorkeudeksi asetettiin 0,07 metriä puun syntypisteestä ja maksimiksi 0,5 metriä. Mikäli kannonkorkeus oli rajojen ulkopuolella, valittiin tapauskohtaisesti kannolle joko minimi- tai maksimi-arvo ja suoritettiin uusintasovitus. Myös minimi- ja maksimiläpimittaa tarkasteltiin rajoittein. Rinnankorkeusläpimitta määritettiin lausekkeella:

$$(3.9) \quad x = 1 - \frac{(1,3 + h_{k2} - h_{k1})}{h_{k2}}$$

, jossa h_{k1} tarkoittaa kannonkorkeuden alkuarvoa ja h_{k2} ensimmäisen vaiheen kannonkorkeutta. STM-tiedon suodatus on tässä vaiheessa mahdollista, mikäli sitä ei ole tehty jo hakkuukoneessa. Suodatettu tieto tarkoittaa ehtoa, joka estää rungon läpimittojen kasvamisen edetessä kannosta kohti latvaa, eli rungon muoto on monotoninen. Suodatus tapahtuu 25 millimetrin toleranssilla.

Toisessa vaiheessa määritetään korjauspolynomi, jolla mahdollistetaan puulajikohtaisesti runkokäyrän parempi sopivuus rungon todellisiin läpimittoihin kolmen ensimmäisen b -parametrin osalta. Kolmannessa vaiheessa rungoille lasketaan korkeuden estimoidut minimi- ja maksimi-arvot, jotka tuotetaan rinnankorkeusläpimitan avulla. Tässä vaiheessa toistetaan runkokäyräsovitusta korjauspolynomi huomioiden. Sovitusvaiheen lopuksi parametrin h sekä h_k vakioidaan.

Neljännessä vaiheessa valitaan sopiva sovituskäyrä riippuen STM-tiedon määrästä. Jos tietoa on vähemmän kuin kolmannes rungon kokonaispituudesta, käytetään kantoläpimitan ja kolmen ensimmäisen b :n parametrin sijaan vain kantoläpimittaa ja ensimmäistä b :n parametria. Tällä estetään käyrän hallitsematon käyttäytyminen.



Kuva 2. Havainnekuva runkokäyrän muodostamisesta STM-tiedon perusteella. Käyrä pyrkii mukailemaan todellisia läpimittoja mahdollisimman tarkasti. Vaaka-akselilla kuvattu suhteellinen pituus latvasta kantaan päin ja pystyakselilla läpimitta millimetreissä.

Runkojen tilavuudet laskettiin integroimalla latvasta tyveen päin. Tavaralajien määrittämistä varten rungosta etsittiin ne kohdat, joissa ideaaliset tukin ja kuidun läpimitat

sijaitsevat. Laskennan tarkastuksessa käytetään alkuperäistä STM-tiedon pituus- ja läpimittatietoa. Jos rungolle ei muodostu tukkipuun minimipituus-läpimittakombinaatioita, merkitään koko puun käyttöosa kuitupuuksi.

3.2 Hilatiedon hankinta

Suomen Metsäkeskuksen (SMK) tuottamaa hilatietoa hyödynnettiin tutkimuksessa. Tätä tutkimusta varten pyydettiin Metsäkeskukselta uusinta laserkeilaustietoa käyttöön Länsi-Uudenmaan alueelta; samalta alueelta, josta hakkuukonetietoa oli jo valmiiksi kerättyä. Metsäkeskus kerää tietoa aluepohjaisella menetelmällä, jossa puustotunnukset ovat yleistetty 16 * 16 m ruudukkotasolle. Hilat sisälsivät tieto- eli attribuuttitaulun, jossa jokaiselle hilalle oli laskettu muun muassa puulajikohtaiset puustotunnukset, kasvupaikkatyyppi sekä maalaji. Puustotunnukset olivat keskimääräinen ikä, hehtaarikohtainen pohjapinta-ala, hehtaarikohtainen runkoluku ja pohjapinta-alalla painotettu puun keskiläpimitta sekä pituus. Laserkeilaustieto on kerätty vuoden 2015 aikana.

Metsäkeskuksen hilatiedon tarkkuuden määrittävät laatukriteerit. Noin 200 000 hehtaarin inventointialueelle sijoitetaan noin 600–800 referenssikoealaa edustavuudeltaan mahdollisimman erilaisille paikoille (Suomen Metsäkeskuksen metsävaratiedon laatuseloste 2016). Hila-aineiston yhden hilaruudun koko on 256 m². Laserkeilauksessa käytetään harvaa pulssitiheyttä, joka on noin 0,5 pulssia per m². Aineiston tavoitetarkkuus on pohjapinta-alalle ± 3 m²/ha, keskiläpimitalle ± 3 cm, keskipituudelle ± 2 m ja keskitilavuudelle ± 20 % 80 % osuudelle aineistosta. Runkoluvun tarkkuudelle ei ole asetettu toleranssia. Puulajin tunnistamistarkkuutta ei ole määritetty prosentuaalisesti, mutta kriteerinä on pääpuulajin määrittäminen, jos kuviolla on selkeästi olemassa sellainen. Puulajien määrittämistarkkuudesta seuraa, etteivät puulajikohtaiset tarkkuudet puustotunnuksissa yllä samaan tarkkuuteen kuin kokonaispuuston tarkkuudessa. Ikä määritetään koealan puustosta, mutta laskennassa sille ei ole määritelty tarkkuuskriteerejä. On huomattu, että suurimmassa osassa tapauksista iän tarkkuus on ± 25 %.

Hilatietoa tuli käsitellä ennen analysointia. Hilatieto toimitettiin kiinteistöittäin erillisinä vektoritiedostoina, jotka tuli ensin yhdistää kokonaiseksi karttatasoksi ja tämän jälkeen kuviot leikattiin hakkuukoneen paikkatiedosta tuotetun kartan avulla. Hakkuukonekuvioista kopioitiin hilasolujen attribuutteihin kuvion tunnistenumero, kuvion

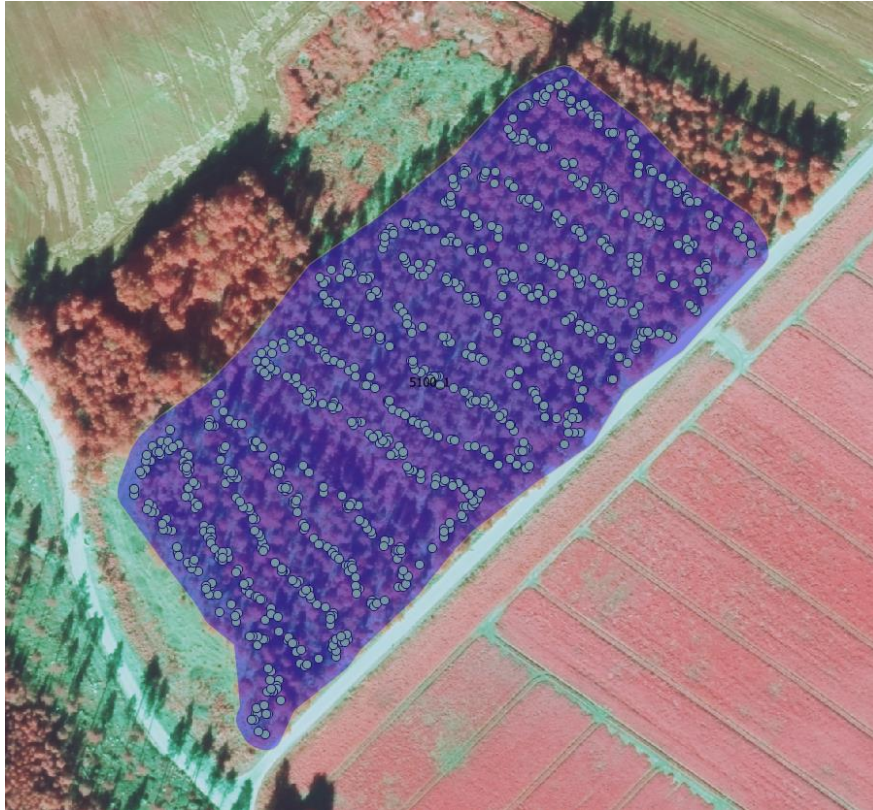
pinta-ala, hakkuutapa sekä kuvion ympärysmitta eli piiri. Kahden viimeisen tunnuksen avulla voitiin laskea kuvion muotoluku, jota voitiin käyttää yhtenä selittäjänä kuvion tarkkuuden suhteen. Attribuuttitaulukosta ilmenee kasvupaikkaluokka, maalaji, oijen kuivatustilanne, ojitusvuosi, puuston kehitysluokka, mahdollinen pääpuulaji, laserkorkeus, eli korkeusjakauman kohta johon mennessä on kertynyt 85% kaikista laserhavainnoista sekä lasertiheys, eli pisteiden osuus yli kahden metrin korkeudella maasta. Puulajikohtaisia puustotunnuksia ovat ikä (T), pohjapinta-ala (G), runkoluku (N), pohjapinta-alalla painotettu keskiläpimitta (DG), pohjapinta-alalla painotettu korkeuden keskiarvo (HG) sekä tilavuus (V), jotka ovat myös ilmoitettu yli puulajiositteiden. Puulajit ovat jaettu kolmeen luokkaan; mänty, kuusi sekä lehtipuut.

3.3 Kuvioiden raja

Hakkuukoneaineiston kuviorajauksessa käytettiin Metsäteholla kehitettyä puoliautomaattista algoritmia. Kaataessa puun, hakkuukone tallentaa dimensioiden lisäksi jokaiselle puulle kaupan sekä tekoalueen tunnusten lisäksi koneen GPS-sijainnin jokaisen kaadon yhteydessä. Sijainnin tarkkuus on muutaman metrin luokkaa. Koska sijaintia ei saada hakkuulaitteesta, ei puun tarkkaa sijaintia voida määrittää. Siksi menetelmässä hakkuukoneen sijainnin ympärille luotiin puskuri, joka kattaa kouran ulottuvuuden. Päätehakkuukuvioilla jokainen puu poistetaan, jolloin puskuri kuvaa riittävän tarkasti hakkuukuvion muodostumista. Tässä tutkimuksessa ei ollut tarkoitus arvioida kuvioinnista muodostuvaa virhettä, mutta oletuksena oli, että sitä esiintyy etenkin pienimmillä kuvioilla.

Prosessin ensimmäinen vaihe oli tiedonkeruu alueella toimineista hakkuukoneista. Hakkuukoneiden tuottamaa StanForD 2010 -tiedonsiirtostandardin mukaista pölkkytietoa piti esikäsitellä, jotta runkojen tiedot saatiin käsiteltävään muotoon (Melkas, henkilökohtainen tiedonanto). Runkojen GPS-pisteille suoritettiin koordinaattimuunnos vanhentuneesta WGS84 koordinaattijärjestelmästä uuteen ETRS FIN35TM -järjestelmään. Aineisto siirrettiin Quantum GIS (Quantum GIS Development Team 2017) paikkatieto-ohjelmaan, jossa puujoukot eroteltiin omiksi kuvioikseen. Tekoaluetunnisteen avulla runkojen ryhmittely oli mahdollista automatisoida. Leimikoille johtavat ajourat oli eroteltava aineistosta. Koneenkuljettajan syöttämät tiedot hakkuutavasta auttoivat erottelemaan aineistosta päätehakkuut. Peruskuviot muodostettiin Delaunay -kolmiointimenetelmän avulla, missä runkojen muodostamien kolmioiden pinta-ala toimi

kuviorajausten rajoitteena (Alliez ym. 2009). Menetelmän tekoon käytettiin pakettia CGAL 4.10 Quantum GIS ohjelmistossa (Quantum GIS Development Team 2017). Puiden pisteiden ympärille muodostettavat ”vyöhykkeet” toimivat kuviorajoina. Peruskuviolla vyöhyke oli kuusi metriä ja ajourassa kaksi metriä. Kuviot voitiin muodostaa siten yhtenäisestä alueesta ja niille voitiin asettaa omat yksilölliset tunnukset.



Kuva 3 Kuviorajauksessa käytetty algoritmi loi kuviorajat. Koska kyseessä oli hakkuukoneen sijainti, piti kouran oletettu etäisyys lisätä vyöhykemäisesti kuvion reunoille. Kuvassa erottuu selvästi hakkuukoneen kulkemat ajourat. (Melkas 2017).

3.4 Motti-mallin kuvaus

Luonnonvarakeskuksen Motti-ohjelmisto on puuston analysointiin ja estimointiin tarkoitettu ohjelmisto (Motti-ohjelmisto 2016). Sitä hyödynnetään pääsääntöisesti puuston kasvun simulointiin erilaisissa skenaarioissa. Skenaariot ja laskennat perustuvat vuosikymmenien aikana kerättyihin maastomittausaineistoihin sekä niistä johdettuihin malleihin.

Tässä tutkimuksessa hyödynnetään ohjelmaa, joka muodostaa puujoukon katkontasimulaattoria varten. Lähtötietona ohjelma hyödyntää puuston keskitunnuksia ja laskennassa käytetään parametrisia menetelmiä. Ohjelmat ovat dynamic link library (*.dll) -muotoisia. Puuston rakenteen kuvaus tapahtuu perinteisenä 1-cm runkolukusarjana tai

vaihtoehtoisesti satunnaisena puujoukkona, joka voidaan viedä suoraan katkontasimulaattoriin. Ohjelma muodostaa puujoukon läpimittajakauman kahden parametrin Weibull-jakauman avulla hyödyntäen parametrin palautusmenetelmää. Menetelmän kuvaus on selitetty tarkemmin Siipilehdon ja Mehtätalon (2013) tutkimuksessa. Satunnaiseen puujoukkoon on puuston keskitunnuksista varioitu läpimittaa ja pituutta, jotta se vastaisi mahdollisimman realistisesti aitoa puujoukkoa.

Parametrin palautusmenetelmä (Parameter Recovery Method = PRM) on vaihtoehto tyypillisesti käytetylle parametrin ennustamismenetelmälle (Parameter Prediction Method = PPM) (Siipilehto ja Mehtätalo 2013). Sen etuna on useamman puustotunnuksen hyödyntäminen Weibull-jakauman ratkaisemisessa runkolukusarjan luomiseksi. Koska ratkaistusta Weibull-jakaumasta voidaan johtaa halutut puustotunnukset, voidaan tiedettyjä puustotunnuksia sijoittaa myös yhtälön toiselle puolelle, jolloin ne ovat osa lähtötietoa. Annetuista puustotunnuksista hyödynnettiin pohjapinta-alalla painotettua läpimittaa (DG) Weibull-jakauman ratkaisussa.

Käytetty kahden parametrin Weibull-funktio on muotoa:

$$(3.10) \quad f(x, b, c) = \frac{c}{b} \left(\frac{x}{b}\right)^{c-1} e^{-\left(\frac{x}{b}\right)^c}$$

, jossa x vastaa koealan puun läpimittaa tai pituutta. Parametrin palautusyhtälöt ovat määritelty neljälle eri läpimitan tunnukselle: läpimitan keskiarvolle D, mediaanille DM, pohjapinta-alalla painotetulle keskiarvolle DG sekä pohjapinta-alalla painotetulle mediaanille DGM. Palautusyhtälöt ovat muotoa

$$(3.11) \quad b\Gamma\left(\frac{1}{c} + 1\right) - D = 0,$$

$$(3.12) \quad b \frac{\Gamma(3/c+1)}{\Gamma(2/c+1)} - DG = 0$$

$$(3.13) \quad b \ln 2^{\frac{1}{c}} - DM = 0$$

$$(3.14) \quad b(F^{-1}\left(0.5; \frac{2}{c} + 1\right))^{\frac{1}{c}} - DGM = 0$$

$$(3.15) \quad b^2\Gamma\left(\frac{2}{c} + 1\right) - DQM^2 = 0$$

, joista viimeinen yhtälö kuvaa ns. toista momenttia. Toinen momentti voidaan laskea tunnuksista G ja N kaavalla $G/(\pi/200^2) * N$.

Mikäli aineistosta puuttuu jokin puustotunnus, ne ennustetaan. Puustotunnusten ennustamisessa käytetään mäntypuulajille Siipilehdon (2011a) malleja, jotka tuottavat tunnukselle parhaan lineaarisen harhattoman ennusteen (Best Linear Unbiased Predictor = BLUP). Muille puulajeille puustotunnusennusteet tehdään erikseen laadittujen mallien

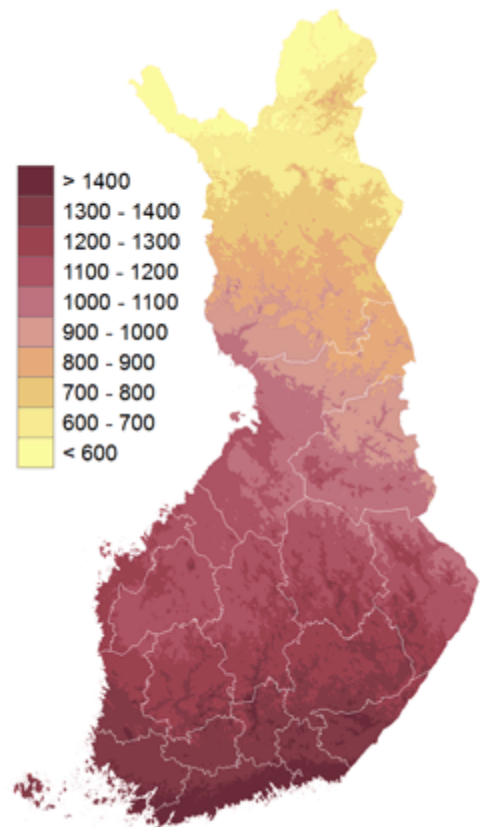
kanssa (Siipilehto, henkilökohtainen tiedonanto). Menetelmää hyödynnetään yleisesti lineaaristen sekamallien satunnaisefektien ennustamisessa. Tarvittavia puustotunnuksia on kahdeksan (G, N, D, DGM, H, HGM, Ddom, Hdom), joista kaksi viimeistä viittaavat valtapuiden läpimittaan sekä pituuteen. Tunnuksilla D, Ddom tai Hdom ei ole varsinaisesti merkitystä laskennassa, mikäli DG, HG sekä N tunnetaan ja nämä tunnuksien ovat keskenään loogisia. Tunnukset G, N ja DG ratkaisevat aineistossa Weibullin jakauman. Tunnuksia DGM ja HGM ei ennusteta, sillä ne ovat todellisuudessa hyvin lähellä tunnuksia DG ja HG (Siipilehto, henkilökohtainen tiedonanto). Sen sijaan jälkimmäisiä tunnuksia käytetään kalibrointimuuttujina tunnusten DGM ja HGM paikalla.

Lähtötiedossa voi joskus olla laskennasta johtuvia virheitä, jonka vuoksi aineistolle voidaan tehdä loogisuustestejä. Jos toisen momentin neliökeskiarvo D_q on yhtäsuuri tai suurempi kuin pohjapinta-alalla painotettu keskipituus DG, voidaan todeta arvojen olevan keskenään epäloogisia, eikä yhtälöä voida ratkaista (Hardy ym. 1988; Siipilehto ja Mehtätalo 2013). Yleensä lähtötiedon virheet johtuvat virheellisistä G tai N tunnuksista ja epäyhtälön jäädessä toteutumatta joudutaan toista näistä tunnuksista muuttamaan käyttäen Nissisen (2002; Siipilehto ja Mehtätalo 2013) puustotunnusmalleja.

3.5 Puujoukon ennustaminen ja laskenta

Hilasolujen attribuuttitaulukkoa käytettiin ohjelman laskennassa. Laskennassa puulajikohtaisia puustotunnuksia olivat ikä, pohjapinta-ala, runkoluku, keskiläpimitta sekä keskipituus. Yhteisiä kuviotunnuksia olivat kasvupaikka, maalaji, ojien kuivatustilanne ja lämpösumma pitkän ajan keskiarvona. Lämpösummana käytettiin vakioarvoa $1293\text{ }^{\circ}\text{Cvrk}$, joka saatiin hila-aineiston tuottajalta. Arvo kuvaa hyvin alueen keskiarvoa (Kuva 4).

Ohjelma tuottaa jokaiselle kuviolle oman puujoukon sekä runkokäyrät. Puujoukko esitetään



Kuva 4. Termisen kasvukauden keskimääräinen lämpösumma vertailukaudella 1981-2010 (Terminen kasvukausi 2017)

taulukkomuodossa, jossa yksi rivi edustaa yhtä puuta. Rivillä on esitetty puunumero, puulaji, läpimitta sekä pituus. Runkokäyrätiedostossa yhdeltä riviltä ilmenee puun numero, puulaji ja tämän jälkeen läpimitta rungonosissa kymmenen senttimetrin välein alkaen 10 cm korkeudelta 300 senttimetriin asti.

Hakkuukonetiedon ja Motti-mallin rungoista muodostettiin kullekin kuviolle puujoukko puulajeittain, jossa jokaiselle puulle oli annettu pituus- ja läpimittaparametri. Runkojen yhteisiä tunnuksia olivat kuviotunnus, kuvion pinta-ala sekä hakkuutapa. Runkojen yksilöllisiä arvoja olivat puulaji, rungon numero, runkoluku, pituus metreinä, läpimitta senttimetreinä sekä tilavuus kuutiometreinä. Kullekin rungolle laskettiin myös pohjapinta-ala. Annetuista tiedoista laskettiin puulajikohtaiset sekä summatut puustotunnukset jokaiselle kuviolle.

Kuviotunnukset luotiin laskemalla puujoukosta summa- ja keskitunnukset. Merkittävimmät summatunnukset olivat pohjapinta-ala hehtaarilla (G , m^2/ha), puiden lukumäärä hehtaarilla (N , kpl/ha) ja tilavuus (m^3/ha). Keskitunnuksia olivat pohjapinta-alalla laskettu läpimitan keskiarvo (DG , cm) sekä pohjapinta-alalla painotetun pituuden keskiarvo (HG , m). Kuviotunnukset laskettiin R-ohjelmistossa (R Core Team 2017). Pohjapinta-alalla painotettuihin tunnuksiin käytettiin painotuskomentoa. Runkojen tilavuus laskettiin syntypisteestä ylöspäin. Motti-mallin tuottamille puille laskettiin tilavuustunnukset käyttäen apuna mallin tuottamaa runkokäyrää, jossa läpimitat olivat ilmoitettu kymmenen sentin välein. Runkojen tilavuudet laskettiin hyödyntäen Huberin kaavaa. Koska runkokäyrä alkoi kymmenen senttimetrin korkeudelta, jouduttiin tyviläpimitta estimoimaan. Tämä tapahtui siten, että kymmenen senttimetrin korkeudelta ilmoitettu läpimitta summattiin kymmenen ja kahdenkymmenen sentin korkeudelta ilmoitetun läpimitan erotuksen kanssa.

Muodostettuja tilavuuksia verrattiin Laasasenahon (1982) tutkimuksessa esitettyyn kahden parametrin tilavuusyhtälöön, jossa tilavuus lasketaan rinnankorkeusläpimitan ja pituuden sekä puulajin avulla. Motti-mallilla tuotetun puuston erotus näiden kahden menetelmän välillä oli pieni. Kun molempien menetelmien puuston kokonaistilavuudet laskettiin yhteen, oli tilavuusyhtälöllä laskettu määrä 2,56 % pienempi kuin runkokäyrältä laskettu tilavuus. Puulajikohtaisesti männyn tilavuus kasvoi 1,05 %, kuusen tilavuus pieneni 3,80 % ja lehtipuiden 1,67 %. Todellisten runkojoukkojen kohdalla tilavuusyhtälöiden aliarvio oli kokonaispuustossa vähemmän, 0,57 %. Männyn tilavuudessa oli 3,99 % yliarvio, kuusen tilavuuden aliarvio oli 1,46 % sekä lehtipuiden kohdalla yliarviota esiintyi 0,82 % runkojen tilavuudet tilavuusyhtälöillä laskettuna.

Vertailu toimi osoituksena siitä, että ilman runkokäyriäkin tilavuus voidaan estimoida kohtuullisen luotettavasti pituuden ja rinnankorkeusläpimitan avulla.

Katkontatuloksista laskettiin kunkin puulajin tavaralajien yhteenlaskettu tilavuus sekä runkojen lukumäärät. Katkontasimulaattorissa käytetty katkontaohje sisälsi tyypillisimmät tukki- ja kuitupuutavaralajien mitat. Se ei sisältänyt mm. pikkutukki- tai muita erikoispuutavaralajeja, joten lopullinen jaottelu voitiin tehdä tukki- ja kuitupuutavaralajeihin. Tukkirunkojen eri laatuluokat yhdistettiin tarkasteluvaiheessa yhdeksi, sillä simulaattori ei sisältänyt laatuparametreja.

3.6 Runkolukusarjojen yhteensopivuustestit

Motti-mallin tuottamia kuvioittaisia runkolukusarjoja verrattiin aitoihin runkolukusarjoihin erilaisilla yhteensopivuustesteillä. Yhteensopivuustestejä käytetään tilastotieteessä nollahypoteesin testaamiseen, onko otosjakauma peräisin tietystä perusjoukosta (Reynolds ym. 1988). Tavanomaisesti yhteensopivuustestit vertaavat otoksen jakaumaa johonkin tunnettuun jakaumaan, kuten normaali- tai Weibull-jakaumaan (Reynolds ym. 1988). Siksi kahden läpimittajakauman samankaltaisuuden arviointiin sopivat huonosti testit, jotka ovat lähtökohtaisesti suunniteltu normaalijakaumaoletuksen testaamiseen. Mikäli testiä käytetään läpimittajakaumien testaamiseen, siinä on oltava mahdollisuus tehdä myös kahden otoksen testi.

Reynolds'n ym. (1988) tutkimuksessa arvioitiin muutamia jakaumien vertailuun soveltuvia yhteensopivuustestejä, kuten Kolmogorov-Smirnov, Anderson-Darling, Cramér von Mises sekä X^2 (Khiin neliötesti). Perinteisten yhteensopivuustestien ongelmana on kuitenkin, että ne antavat yhtä paljon painoa jokaiselle läpimittaluokan frekvenssille, vaikka pieniläpimittaisten puiden merkitys on metsikkökuviolla etenkin taloudellisesta näkökulmasta pienempi. Käyttämällä läpimittaluokan pohjapinta-alaa painotuksessa voidaan suuriläpimittaisten frekvenssien merkitystä kasvattaa lopputuloksessa (Gove ja Patil 1998).

Tutkimukseen valikoituneet yhteensopivuustestit olivat Kolmogorov-Smirnov (K-S) -yhteensopivuustesti sekä virheindeksimenetelmä (Reynolds ym 1988), jota tai sen muunnelmia on käytetty useassa tutkimuksessa (Peuhkurinen ym 2007, Peuhkurinen ym. 2008, Packalén ja Maltamo 2008, Peuhkurinen ym. 2011, Siipilehto ym. 2016).

Läpimittaluokat laskettiin yhden senttimetrin luokissa. Analyysi suoritettiin käyttäen R-ohjelmistoa (R Core Team 2017). Yhteensopivuustestien lähtökohtana on tarkastella, päteekö nollahypoteesi kahden jakauman välillä. Lauseke voidaan esittää seuraavasti:

$$(3.16) \quad H_0: F(x) = F_0(x)$$

, jossa $F(x)$ on kumulatiivinen yksiulotteinen jakauma satunnaisesta muuttujasta X ja $F_0(x)$ tarkoittaa kumulatiivista jakaumaa todellisista arvoista. Mikäli $F_0(x)$ arvot ovat kokonaan määritelty, testaukseen voidaan soveltaa K-S testiä. Kustakin frekvenssistä lasketaan sen suhteellinen osuus koko sarjasta. Arvot ovat kumulatiivisesti sijoitettu välille nolla ja yksi suuruusjärjestyksessä alkaen pienimmästä läpimittaluokasta. Kolmogorov-Smirnov -kahden otoksen yhteensopivuustestin kaava on muotoa:

$$(3.17) \quad D = \max_x |F^*(x) - F_0(x)|$$

, jossa $F_0(x)$ kuvaa nyt todellista jakaumaa ja $F^*(x)$ vastaavasti jakauman muutosta kussakin luokassa. Tarkoituksena on löytää *supremum*, eli suurin erotus yhden luokan sisällä kumulatiivisessa jakaumassa. Kaavan itseisarvon on tarkoitus viitata siihen, että kyseessä on suurin erotus luokkien välillä riippumatta siitä, onko estimaatti havaittua korkeampi tai matalampi kyseisessä luokassa.

Tässä tutkimuksessa verrattiin Weibull-jakaumasta laskettuja läpimittajoukkoja aitoihin läpimittajoukkoihin. Molemmat jakaumat ovat siten empiirisiä. Sokal ja Rohlf (2012) toteavat K-S testin soveltuvan hyvin jatkuvien jakaumien tutkimiseen, mutta diskreettien jakaumien tutkimiseen testi on vanhoillinen. Toisin sanoen sen soveltuvuudesta tähän tutkimusaineistoon löytyy ristiriitaisuuksia. Testin hyvyttä voidaan tutkia vertailemalla tuloksia muihin yhteensopivuustesteihin. Mitä tulee Anderson-Darling ja Cramér von Mises -testeihin, tutkivat ne koko jakauman muutosta. Anderson-Darling -testi painottaa enemmän tuloksessa jakauman ääripäitä kuin Cramér von Mises, mikä on hyvä asia jakauman suuriläpimittaisten puiden kohdalla, mutta toisaalta huono silloin kun pieniläpimittaisten puiden frekvensseille annetaan suuri arvo. Testien käyttöä hankaloitti myös se, ettei kahden otoksen kriittisten arvojen laskukaavoja löytynyt tutkimuskirjallisuutta etsiessä. Pelkkä testistä saatu arvo ei auta tulosten tulkitsemisessa. Testejä varten voidaan pieniläpimittaiset, esimerkiksi läpimitaltaan alle 10 cm, puut jättää pois laskennasta, mutta se ei poista ongelmaa painotuksen suhteen.

K-S testi tutkii kahden empiirisen kumulatiivisen jakauman eroa ja etsii suurinta yhden luokan sisällä olevaa määräeroa D . Kahden muuttujan K-S testissä nollahypoteesin merkitsevyystason laskeminen ei ole tarkasti laskettavissa, joten sen sijaan voidaan käyttää

populaation koosta laskettavaa D :n raja-arvoa. Tämä raja-arvo määritetään lausekkeella (Sokal ja Rohlf 2012):

$$(3.18) D_{lim} = \sqrt{-\frac{\ln \frac{\alpha}{2}}{2n}}$$

, jossa α tarkoittaa haluttua merkitsevyystasoa ja n havaintojen lukumäärää. Tässä tutkimuksessa käytetty merkitsevyystaso oli 5 %, joka on yleisyytensä vuoksi helpoiten tulkittavissa. Tuloksista saadaan vertailukelpoisia, kun arvo D jaetaan arvolla D_{lim} (Siipilehto ym. 2016, Tham 1988). Siipilehto ym. (2016) esittivät, että mikäli saatu osamäärä on alle yhden, on jakauma läpäissyt K-S testin (ym 2016). Vastaavasti Tham'n (1988) tutkimuksessa hyvänä tuloksena pidettiin lukuarvoltaan alle neljän osamäärää läpimittajakaumassa.

Reynolds'n ym. (1988) tutkimuksessa esiteltyä virheindeksiä (error index) käytettiin arvioimaan koko kuvion jakauman vaihtelua. Sen kaava on muotoa

$$(3.19) e = \sum_{i=1}^k w_i |f_i - \hat{f}_i|$$

, jossa k on läpimittaluokkien lukumäärä, f_i on runkojen havaittu määrä ja \hat{f}_i on runkojen estimoitu määrä läpimittaluokassa i , w_i on paino luokassa i . Painotuksena käytettiin läpimittaluokan pohjapinta-alaa.

Edellä mainitut tilastolliset menetelmät arvioivat runkolukusarjojen yhteensopivuutta, mutta eivät kuvaa runkolukusarjan muotoa, jolla voi olla merkitystä puuston ennustamistarkkuuden kanssa. Runkolukusarjat piirrettiin pohjapinta-alalla painotettuina puulajeittain sekä puulajit yhteenlaskettuina. Kuvaajista voitiin tutkia syitä tilastollisten menetelmien antamiin arvoihin sekä tutkia myös sitä, milloin runkoluvut vastasivat toisiaan ja milloin ei. Kuvaajia tutkiessa huomattiin, että selkeästi yksihiippuiset puujoukot kuvautuivat parhaiten ja pääsyy käyrän muotoon oli kuvion pinta-ala; kun puustoa on vähän, frekvenssit vaihtelevat suuresti läpimittaluokkien välillä. Pinta-alan ja runkoluvun kasvaessa läpimittajakaumat tasoittuivat, jolloin luokkien välillä ei ollut suurta vaihtelua. Pienillä kuvioilla läpimittajakaumien vaihtelu oli sen sijaan tavanomaisempaa.

Kuvioiden todellisia läpimittajakaumia haluttiin arvioida suhteessa katkonnan tarkkuuteen. Koska yksihiippuisilla jakaumilla jakaumat vastasivat yleensä toisiaan, haluttiin asiaa tutkia enemmän tilastollisesti. Jokainen kuvio arvotettiin silmämääräisesti kolmeen eri luokkaan. Luokka 1 sisälsi vain selkeästi yksihiippuiset läpimittajakaumat. Toisen luokan läpimittajakaumat olivat selkeästi kaksihiippuisia, tai selkeää huippua ei voitu erottaa. Kolmanteen luokkaan kuuluivat loput, eli kuviot, joiden todellisissa

läpimittajakaumissa ei ollut selkeää säännönmukaisuutta. Havainnointi oli täysin yhden havainnoitsijan subjektiivinen näkemys. Siten virhettä ei aiheutunut usean eri havaitsijan satunnaisefektistä, mutta menetelmää ei toisaalta sen subjektiivisuudesta johtuen voida pitää kovin luotettavana. Etenkin pienten kuvioiden läpimittajakaumissa yksihuippuisuuden havaitseminen on haasteellista, sillä luokkien välinen hajontaa esiintyi runsaasti. Pienet vaihtelut eri luokkien välillä jätettiin huomioimatta kyseisillä kuvioilla ja ne tulkittiin helpommin yksihuippuisiksi.

Motti-mallin käyttämä Weibull- jakauma on yksihuippuinen (Mehtätalo ym. 2007, Siipilehto ja Mehtätalo 2013), mikä tukisi hyvin päätelmää, että malli tuottaisi tarkasti puustot kuvioille, joissa toteutuu läpimittajakauman yksihuippuisuus. Koska jakaumat tuotetaan kullekin hilalle erikseen, on yhdelle hilalle tuotettava puulajikohtainen jakauma aina yksihuippuinen. Malli voi tuottaa eri jakaumat eri vaiheiden puustoille, mikä on tarpeen etenkin, jos kuviolla on runsaasti kuusialikasvosta. Oletuksena siitä huolimatta oli, että yksihuippuiset puulajijakaumat ovat myös helpoiten ennustettavissa.

3.7 Keskitunnusten sekä katkonnan tulosten vertailu

Keskitunnukset laskettiin kolmesta eri lähteestä ja tuloksia verrattiin keskenään. Sekä hakkuukonetulosteesta poimituille että hilatiedosta Motti-mallin tuottamille runkojoukoille laskettiin puustotunnukset G, N, DG, HG, sekä V. Myös hilatiedon hilojen keskitunnuksista laskettiin kuviokohtaiset keskiarvot. Kaikkia keskitunnuksia verrattiin toisiinsa. Taustaoletuksena oli, että hilatiedon ja Motti-mallin keskitunnukset vastaisivat pitkälle toisiaan; tämä todistaisi mallin toimivuuden ja kyvyn muodostaa oikeanmukainen runkojoukko annetuista keskitunnuksista. Toteutuneisiin arvoihin vertailemalla voitiin nähdä ennen kaikkea hilatiedon paikkansapitävyys sekä Motti-mallin asettuminen näiden arvojen väliin. Metsäkeskuksen aineistosta tiedettiin, ettei esimerkiksi runkoluvulle tai puulajiosuuksille ole annettu selkeitä laatuksiteerejä. Siten on mahdollista, etteivät puustotunnukset ole aina täsmällisiä. Mallin käyttäytyminen puustotunnuksiltaan epäselvissä kuvioissa haluttiin selvittää vertailemalla näitä kahta estimaattia toteutuneisiin arvoihin. Parhaassa tilanteessa Motti-mallin keskitunnukset sijoittuisivat jonnekin hilatiedon ja oikean tiedon välille.

Tuloksia vertailtiin harhan (bias) sekä keskineliövirheen neliöjuuren ($RMSE = \text{Root Mean Squared Error}$) avulla.

Keskineliövirheen neliöjuuri laskettiin kaavalla,

$$(3.20) \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

ja harha

$$(3.21) \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$$

, jossa n on kuvioden lukumäärä, y_i on havaittu arvo kuvion muuttujasta y sekä \hat{y}_i tämän estimoitu arvo. Harha ilmaisee tulosten keskimääräistä poikkeavuutta todellisista arvoista alkuperäisissä yksiköissä. Jos estimaatti säännönmukaisesti yli- tai aliarvioi todellisia tuloksia, ilmaisee harha tämän tyyppisen poikkeavuuden. Keskineliövirheen neliöjuuri ilmaisee virheen absoluuttisen suuruuden alkuperäisissä yksiköissä, mikä tarkoittaa, etteivät yli- tai aliarviot eliminoi toisiaan. Harhasta ja virheestä laskettiin myös suhteelliset arvot, missä saatu arvo jaettiin havaittujen arvojen keskiarvoilla.

Puulajikohtaisesti keskitunnusten DG ja HG harhan ja keskivirheen laskentaan käytettiin vain havaintoja, joista löytyi sekä havaittu että estimoitu arvo. Jos havainto puulajin rungoista puuttui, ei keskipituutta tai läpimittaa voitu määrittää. Laskennasta poistettiin ne arvot, joista oli vain yksi havainto. Useimmin puuttuva arvo oli havaittu, sillä hilan kuviotiedoissa jokainen kuvio sisälsi ainakin hieman jokaista puulajiluokkaa. Sen sijaan summatunnusten laskennasta G, N sekä V nolla-arvot olivat mukana. Hilatiedon keskitunnukset laskettiin kuviokohtaisesti siten, että valmista puustotunnuksista laskettiin keskiarvot painottamalla arvoja hilan koolla. Kuviorajauksessa leikkautuneet hilat saivat siten laskennassa pienemmän painotuksen sekä niiden tuottamaa häiriötä voitiin minimoida. Pohjapinta-alalla painotettuja keskitunnuksia DG sekä HG painotettiin lisäksi hilan pohjapinta-alalla, jotta tulokset olisivat totuudenmukaisempia. Puulajikohtaisissa arvoissa käytettiin painotukseen puulajikohtaista pohjapinta-alaa.

Katkongan tulosten tarkastelu painottui etenkin puulajikohtaisten tilavuuksien, tukkipuuprosenttien sekä puutavaralajien keskipäreyksien tarkasteluun. Vaihtoehtoisesti tuloksia olisi voitu tarkastella jakauma-asteen (Bergstrand 1989; Kivinen 2004) avulla. Jakauma-astetta hyödynnetään tutkittaessa tavoitejakauman suhdetta toteutuneen puutavaralajijakauman jokaiseen pituus- ja läpimittaluokkaan kullekin tavaralajille (Kivinen 2004). Jakauma-asteen toimivuus yhteensopivuustestinä on erinomainen, sillä sen tulos on indeksinumero nollan ja yhden väliltä. Tässä tutkimuksessa sen käyttö olisi edellyttänyt, että todellisista rungoista toteutuneen katkonnan jakauma-aste olisi ollut 1, eli kyseinen katkontatulos olisi noudattanut tavoitejakaumaa. Tutkimuksen todellisten

pölkkyjen tavoitejakauman mediaani oli 66 %. Myös puutavaralajien määriä olisi pitänyt käsitellä suhteellisina, sillä jakauma-asteen toiminta edellyttää, että eri pituus- läpimittaluokkien yhteenlaskettu pölkky määrä on sama. Tässä prosessissa katoaa informaatiota määräeroista, eli vaikka puuston runkojen läpimitta- ja pituusjakaumat olisivat olleet samanlaiset, pölkkyjen kokonaislukumäärän muutosta tulos ei olisi pystynyt ilmaisemaan. Jakauma-asteen käytöstä luovuttiin tässä tutkimuksessa sen vähäisen lisäarvon ja suuren työmäärän vuoksi.

Jakauma-asteen sijaan perustui tulosten tarkastelu katkottujen tukkirunkojen kokonaistilavuuksien sekä keskijäreysien tutkimiseen. Sulkemalla kuiturungot pois tarkastelusta saadaan parempi käsitys puuston samankaltaisuudesta simuloidun ja havaitun välillä. Väitteen voi perustella siten, että katkonnan kokonaistilavuus saattaa hyvinkin olla simuloinneissa samanlainen, mikäli runkojen dimensioiden muutos kompensoituu lukumäärän muutoksella. Koska tukkipuiden katkontaa rajoittaa rungon minimiläpimitta, ei pieniläpimittaisemmista puista valmistu yhtä paljoa tukkirunkoja. Vaikka menetelmä ei kerro eksaktia totuutta tukkipölkkyjen yhdenmukaisuudesta, on se silti nopea ja riittävän tehokas menetelmä tätä aineistoa varten.

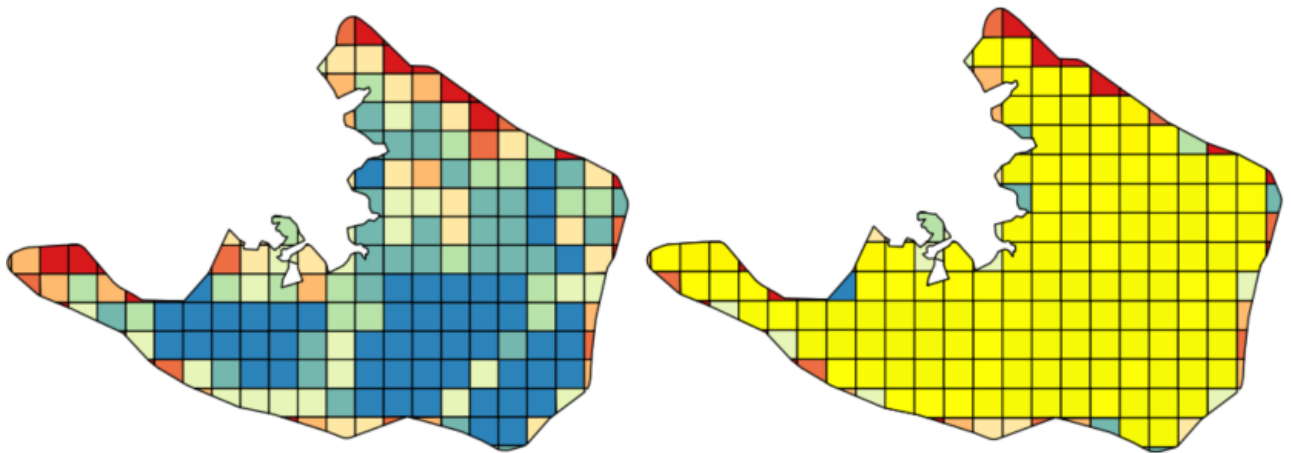
Puuston samankaltaisuus laskettiin tukkirunkojen kokonaistilavuudesta. Symmetrisen keskiarvon absoluuttinen prosentuaalinen virhe (MAPE, Mean absolute percentage error) on keino mitata ennustuksen tarkkuutta suhteellisenä osuutena havaitun ja estimoidun arvon keskiarvosta (Armstrong 19785; Flores 1985).

$$(3.22) \sum_{i=1}^n |(X_i - F_i)| / [0,5 * (X_i + F_i)] / n * 100$$

Toisin kuin perinteisesti laskettu absoluuttinen prosentuaalinen virhe, jossa virhe jaetaan pelkällä havaitulla arvolla, voidaan lopputulos pakottaa nollan ja yhden välille. MAPE ei yliarvioi virheen määrää silloin, kun alkuperäinen arvo on pieni ja estimaatti suuri (Tofallis 2015). Tämä tasapainottaa tilanteita, joissa havainnoitu pölkkyjen lukumäärä on nolla, mutta estimaatti hyvin korkea (Tofallis 2015, Flores 1985). Koska jakajana toimii havaitun ja estimaatin summa, kuvaa ero paremmin todellista eroa ja se on myös laskettavissa. Tässä tutkimuksessa tukkipölkkyjen tilavuuksien puulajikohtaisesta MAPE:n arvosta laskettiin käänteisluku, mikä kuvastaa simuloinnin tarkkuutta havainnoituun. Männyn suuren tilavuuden virheen vuoksi arvioinnissa käytettiin puulajeilla painottamattomia lukuja, eli tukkipölkkyjen kokonaistilavuutta.

Katkonnan tulosten tarkastelussa tutkittiin reunahilojen vaikutusta katkonnan tuloksiin. Puujoukot muodostettiin simulaattorille kahdella eri tavalla hilatiedosta; alkuperäisessä

mukana olivat kuvioiden kaikki hilat ja vertailuaineistossa kuvioiden reunahilat, joiden pinta-ala oli alle 75 % hilan kokonaispinta-alasta (256 m^2) jätettiin pois laskennasta. Tällä menettelyllä haluttiin eliminoida häiriö, jota esiintyi kuvioiden reunahiloilla. Eri kehitysluokan ja/tai puulajia edustavat vieruskuviot, tiestöt, vesistöt, pellot sekä muut maantieteelliset tekijät luovat häiriötä reunahilojen puustotunnuksiin ja siten kuviokohtaisia tunnuksia laskettaessa myös kuvioiden keskitunnuksiin. Automaattisen kuviorajauksen myötä jotkin kuviorajat eivät muodostuneet selkeiksi suoriksi linjoiksi. Suodatus auttoi myös näissä tapauksissa, sillä keskitunnukset otettiin vain hiloilta, joiden puusto oli todennäköisimmin poistettu (Kuva 5).



Kuva 5. Kuvion reunahilojen tilavuus kuvassa on alle $75 \text{ m}^3/\text{ha}$, kun keskellä tilavuus on noin $250\text{--}350 \text{ m}^3/\text{ha}$. Automaattinen kuviorajaus ei aina onnistu tuottamaan suoria linjoja, kuten kuvion vasemmassa reunassa. Mallissa, joka sisälsi reunahilojen suodatuksen, laskentaan sisällytettiin vain keltaisella värjätty hila (oikea kuvio). Kuvion koko $3,7 \text{ ha}$.

Katkonnan tuloksista saadut tavaralajien tilavuudet sekä pölkkyjen lukumäärät skaalattiin tämän jälkeen suhdeluvulla, joka laskettiin siten, että kuvion alkuperäinen pinta-ala jaettiin laskentaan käytettyjen hilojen pinta-alojen summalla. Koska joukossa oli useita pieniä kuvioita, joiden pinta-ala oli puolen ja kahden hehtaarin välillä, uskottiin reunahilojen suodatuksella olevan suuri merkitys näiden kuvioiden tilavuustunnuksiin.

3.8 Korrelaatiokerrointen laskenta

Laajan aineiston vuoksi voitiin tutkia, kuinka eri puustotunnukset selittävät analyysien tuottamaa virhettä. Peuhkurinen ym. (2011) tutkivat eri puustoselittäjien suhdetta selitettäviin muuttujiin, joita olivat Reynolds:n virheideksi pohjapinta-alalla painotettuna sekä tilavuuden virhe läpimitaltaan yli 15 cm puissa. Suhdetta tutkittiin Pearsonin korrelaatiokerrointen avulla. Selittäviä muuttujia olivat puustotunnuksista runkoluku,

pohjapinta-ala, latvuspeittävyys, läpimitan ja pituuden aritmeettiset keskiarvot ja näiden variaatiokertoimet sekä puiden hajontaindeksi. Vastaavanlaista lähestymistapaa käytettiin tässäkin tutkimuksessa merkitsevien selittäjien löytämiseksi. Selittäjien hyvyttä arvioitiin 5 % merkitsevyystasolla. Selittäjien vaikutusta tutkimalla voidaan arvioida, kuinka eri puustotunnukset vaikuttavat tarkkuuteen.

Tutkimuksessa selitettävät ja selittävät muuttujat muodostettiin mukaillen Peuhkurisen ym. (2011) tutkimusta. Selittäviä muuttujia olivat puustotunnuksista pohjapinta-ala, runkoluku, aritmeettinen keskipituus, pituuden ja läpimitan varianssikertoimet, kasvupaikkatyyppi sekä maalaji. Selitettäviä muuttujia olivat virheindeksi pohjapinta-alalla painotettuna sekä katkonnan tuloksesta poimittujen tukkipuiden kokonaistilavuuden tarkkuus. Tarkkuus määriteltiin siten, että laskettiin ensin estimoinnin ja todellisen tuloksen absoluuttinen erotus, joka jaettiin näiden lukujen summalla. Luku saatiin siten pysymään nollan ja yhden välillä. Kun tämä luku vähennettiin yhdestä, saatiin selville suhteellinen tarkkuus. Tulokset sovitettiin R-ohjelman lineaariseen malliin ja tutkittiin korrelaatiokerrointen suuruuksia sekä p-arvojen merkitsevyystasoja.

4 TULOKSET

4.1 Mallirajoitteet

Motti-ohjelman rajoitteet sekä loogisuustarkastukset rajoittivat puujoukkojen laskentaa kaikille hiloille. Ohjelma ei pystynyt ennustamaan puujoukkoa hiloille, joissa yhdenkin puulajin läpimitta oli alle kahdeksan senttimetriä. Läpimitan arvon tuli siten olla joko nolla tai vähintään kahdeksan senttimetriä. Kuvioita tarkastellessa tämä rajoite ei merkittävästi haitannut laskentaa, sillä suodatus poisti hiloja yhteensä 3,27 hehtaaria. Hilojen pinta-alan osuus oli 0,32 %. Hilojen suodatuksella ei täten katsottu olevan suurta merkitystä laskentaan.

Hilasoluja, jotka eivät täyttäneet loogisuustarkastusta $DA < Dq < DG$, jouduttiin muuttamaan laskentaa varten. Kyseisissä tapauksissa ohjelma ei jätä Weibull-käyrää ratkaisematta, vaan muuttaa lähtöaineiston tunnuksia. Kun Dq arvo oli liian pieni, muutos tehtiin runkolukuun ja kun Dq oli liian suuri, muutos tehtiin pohjapinta-alaan (Siipilehto, henkilökohtainen tiedonanto). Tarkastellessa kunkin hilan puulajikohtaisia

puustotunnuksia, löytyi yhteensä 159 havaintoa, jotka eivät läpäisseet loogisuustarkastusta. Puulajikohtaisesti suurin osuus oli männyllä, 67 havaintoa. Koko aineistosta tapausten osuus oli 0,3 %, joten epäloogisuuksista johtuvia puustotunnuskorjauksia ei voida pitää merkittävänä.

Motti-ohjelma laskee yhdelle hilalle aina kokonaisen puujoukon. Tällöin tulisi yhden hilan puulajiositteen hehtaarikohtaisen runkoluvun olla vähintään $(10000 \text{ m}^2 / 256 \text{ m}^2 \approx 39)$ 39 kpl/ha, joka vastaa yhtä runkoa hilalla (Siipilehto, henkilökohtainen tiedonanto). Mikäli luku on alle tämän, hilalle ei muodostu yhtään puuta. Sirpaloituneiden hilojen kohdalla mahdollisuus tyhjään puujoukkoon on suurempi, sillä hilan koon pienetessä tulisi hehtaarikohtaisen runkoluvun kasvaa samassa suhteessa. Puulajikohtaisesti havaintoja, joissa runkoluku oli yli 0, mutta alle 39, toteutui männyllä 5,1 %, kuusella 2,0 % sekä lehtipuilla 10,9 % hila-aineistosta. Tämän voidaan olettaa olevan yksi tekijä puujoukon aliarviointiin.

Motti-malli ei tuottanut runkokäyriä puille, joiden pituusdimensiot ovat liian suuria. Mikäli rungon pituus on yli 40 metriä, ohjelma ei tuota runkokäyriä, eli läpimittoja kymmenen sentin välein, kyseisille puille, vaikka pituus- ja läpimittatunnukset löytyivät puujoukkotaulukosta. Tämä aiheuttaa pientä vääristymää kuvion keskitunnuksiin ja simulointituloksiin. Tässä aineistossa edellä mainittuja runkoja oli yhteensä seitsemän kappaletta, joista kolme sijaitsi yhdellä kuviolla. Laskentoihin ne eivät siten vaikuttaneet merkittävästi.

4.2 Kuviorajausten ongelmia

Kuvioiden laskennassa ongelmia tuottivat virheet keräysmetodeissa. Puiden ja sitä kautta kuviorajojen muodostuksessa käytettiin apuna hakkuukoneen tuotantotietoja. Merkittävin virhetekijä on toisaalta kuljettaja, joka voi merkitä korjuun tekoalueen tiedot syystä tai toisesta väärin. Syinä voivat olla joko inhimillinen erehdys, huolimattomuus tai kuvioiden suhteellisen pienuuden vuoksi tekoaluetta ei viitsitä muuttaa. Virheet paljastuivat pääosin simulointivaiheessa, kun muutamien kuvioiden hilatiedon ja hakkuukonetiedon väliset puustotunnukset eivät vastanneet lähellekään toisiaan; muodostettujen tavaralajien tilavuudet poikkesivat toisistaan hyvin vahvasti. Tuoreiden ilmakuvien perusteella tehtiin silmämääräinen tulkinta hakkuutavan paikkansapitävyydestä ja luokiteltiin virheelliset kuviot uudelleen.

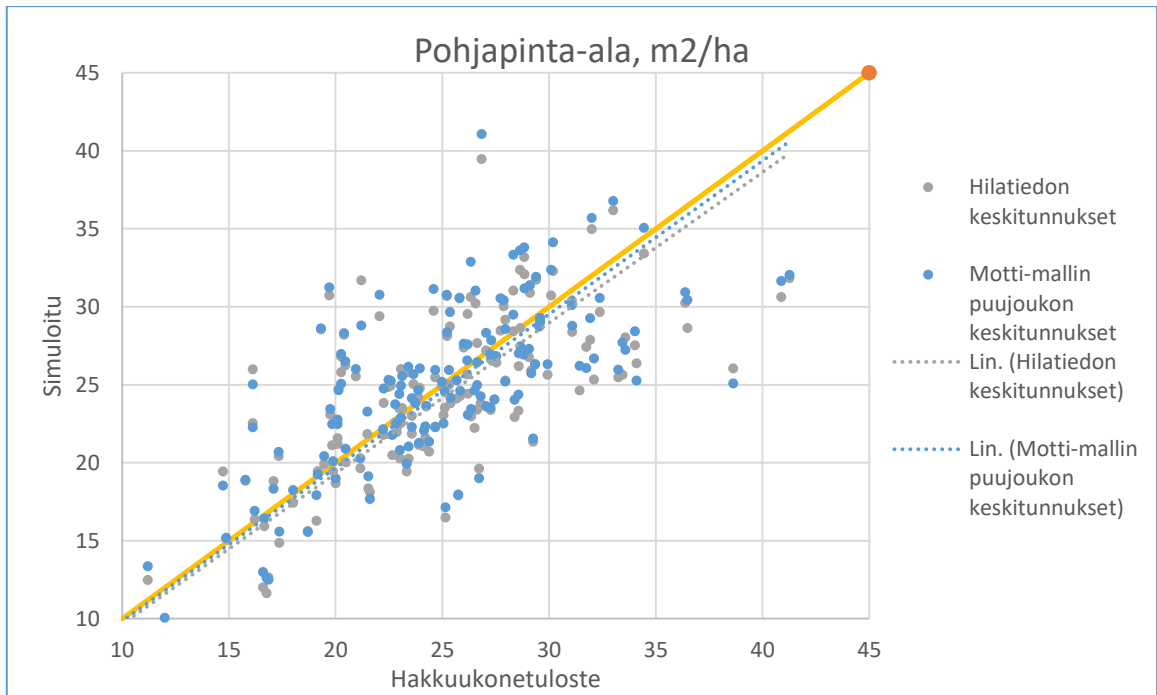
Muita rajaukseen liittyviä ongelmia olivat puiden paikantamiseen liittyvät ongelmat. Yleisin virhetekijä on puun sijainnin puutteellisuus, sillä hakkuupään ja hakkuukoneen välimatka voi olla jopa 10 metriä. Hakkuupään ohjaukseen liittyy toinen ongelma; kun hakkuukone kaataa rungon, sen hakkuupää kääntyy noin 90 astetta ja jää ”kaatoasentoon” rungon prosessoinnin ajaksi. Rungon tiedot tallentuvat hakkuukoneen muistiin siinä vaiheessa, kun hakkuupää nostetaan takaisin pystyasentoon seuraavaa kaatoa varten. Jos hakkuukoneenkuljettaja ei nosta hakkuupäätä pystyasentoon välittömästi ja ajaa koneen pois työalueelta, kaadosta tallentuu väärä koordinaatti järjestelmään. Jos koneenkuljettaja ajaa usein työmaalta pois esimerkiksi huoltamaan konetta tai vuoronvaihtoon ilman hakkuupään nostoa, tallentuu kuviolta useita virheellisiä pisteitä. Näiden suodattaminen automaattisesti on hyvin haastavaa, ja tässä aineistossa suodatus tehtiin manuaalisesti.

4.3 Keskitunnusten tarkastelua päätehakkuukuvioilla

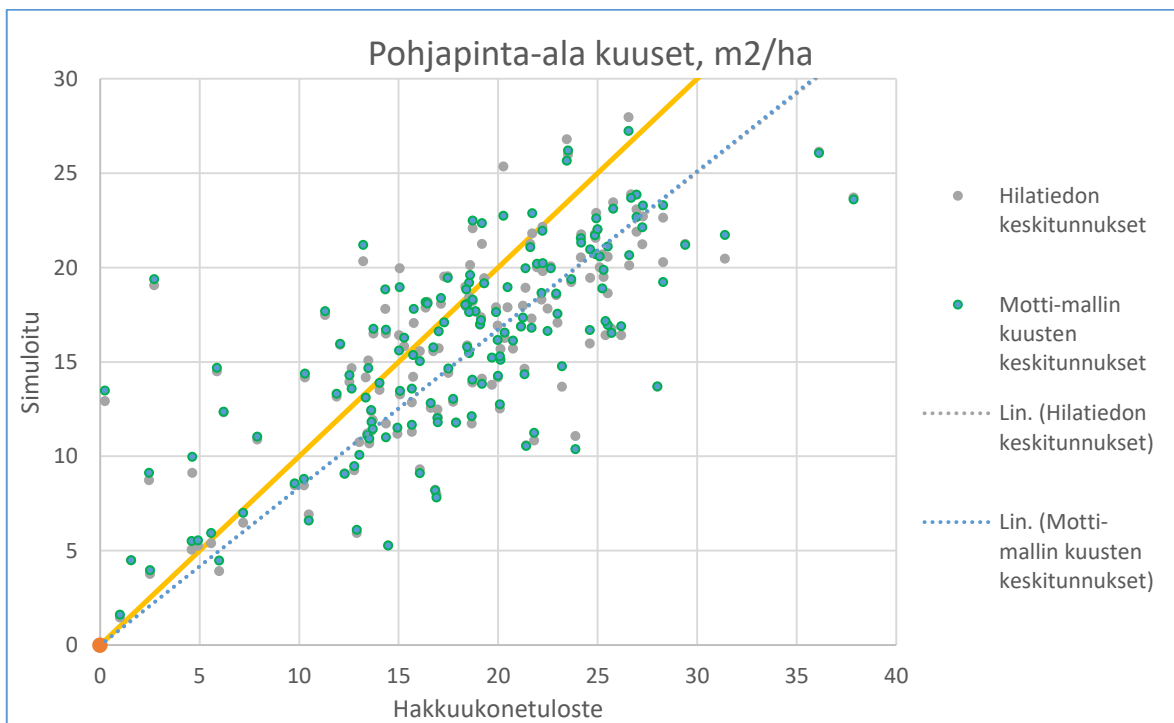
4.3.1 Pohjapinta-ala

Simulointi käsitti yhteensä 150 päätehakkuukuvioita, joiden pinta-ala oli yhteensä 334 hehtaaria. Keskitunnukset laskettiin kahdella tavalla; kaikki hilat mukaan ottaen (Taulukko 2) sekä reunahilojen suodatuksella (Taulukko 3). Koska keskitunnukset olivat hyvin samankaltaisia reunahilojen suodatuksen jälkeen, ei erillisten korrelaatiokuvaajien piirtäminen olisi tuonut merkittävää lisäarvoa laskentaan.

Pohjapinta-alan virheiden tunnuksia tarkastellessa huomataan, kuinka hyvin mallin tuottama puujoukko noudattaa hilatiedon keskitunnuksia (Taulukko 2). Aidon puujoukon ja Motti-mallin tuottaman puujoukon välinen harha on 0%, mutta sen sijaan RMSE 18%. Myös hakkuukonetiedon ja hilatiedon välinen RMSE on 18 %. Pohjapinta-alan ennusteen ei voida olettaa olevan kovin tarkka. Puulajikohtaisesti kuusen pohjapinta-alan estimointi on onnistunut heikoiten. Hilatiedon ja Motti-mallin välinen harha on 11 % sekä virhe 15 %. Erotus ei kuitenkaan ilmene verrattaessa molempia keskitunnuksia hakkuukonetietoon; virhe ja harha ovat lähestulkoon samat. Pistekaaviosta (Kuva 6) voi havaita kuvioiden välillä esiintyvän runsaasti hajontaa. Kuusilla hajontaa esiintyi puulajeista runsaimmin. Kaaviossa on esitetty sekä hilatiedosta lasketut keskitunnukset että Motti-mallin kuvioiden keskitunnukset. Samalle y-akselin arvolle asettuvat hilatiedon ja Motti-mallin pisteet ovat arvoja samalta kuviolta.



Kuva 6. Pohjapinta-alan suhde hakkuukonetulosteeseen ja estimointien välillä. Hilatiedon ja Motti-mallin tunnuksissa on havaittavissa paljon hajontaa suhteessa hakkuukonetulosteesta laskettuihin keskitunnuksiin.



Kuva 7. Pohjapinta-alan suhde oli puulajeista kuusella heikointa.

Taulukko 2. Puustotunnusten virheet ja harhat yksiköin ja osuuksin päätehakkuukuvioilla, jossa G tarkoittaa pohjapinta-alaa, N runkolukua, DG pohjapinta-alalla painotettu keskiläpimitta, HG pohjapinta-alalla painotettu pituus sekä V tilavuus hehtaarilla. Toteutuneet arvot laskettu hakkuukonetulosteesta, Motti-mallilla luodut hilatiedosta sekä hilatiedon alkuperäiset keskitunnukset, n=150.

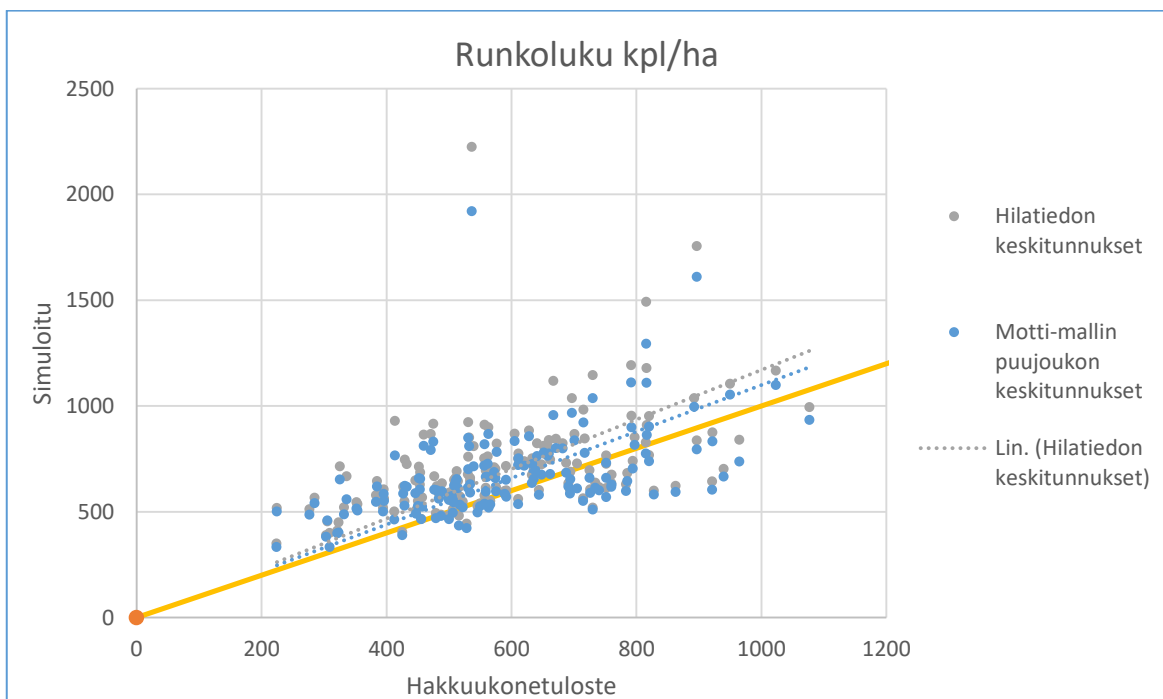
		G, m2/ha				N, kpl/ha				DG, cm				HG, m				V, m3/ha				Puulajiosuudet, %			
		KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	MÄNNYT	KUSET	LEHTIP.	
Hakkuukonetieto ja Motti-malli	harha	-0.1	-2.1	2.2	-0.2	-80.8	-52.1	-3.4	-25.3	1.4	0.5	1.8	1.0	0.8	0.7	0.5	0.9	0.4	-20.7	22.0	-1.0	harha	-8.4	8.3	0.1
	harha-%	-0.3	-84.4	12.2	-3.6	-13.7	-125.1	-0.8	-19.0	4.9	1.6	6.4	4.3	3.4	3.1	2.2	4.2	0.1	-79.9	11.3	-2.2				
	Rmse	4.5	4.0	5.2	2.2	203.4	76.7	162.3	81.1	3.0	6.3	3.6	4.0	1.2	3.1	1.5	2.1	45.2	41.8	54.9	22.7	Rmse	16.2	18.3	8.2
	Rmse-%	17.9	160.6	28.6	50.9	34.4	183.9	39.0	60.7	10.3	20.1	12.6	16.5	5.2	13.5	6.5	9.2	17.1	161.2	28.3	51.9				
Hakkuukonetieto ja hilatiedon keskitunnukset	harha	0.4	-2.0	2.3	0.1	-125.1	-56.8	-27.5	-41.2	1.0	-0.1	1.5	0.8	0.5	0.1	0.4	0.3	6.9	-11.2	23.7	-5.6	harha	-5.1	7.1	-2.0
	harha-%	1.6	-77.4	12.6	1.4	-21.2	-136.4	-6.6	-30.9	3.6	-0.3	5.2	3.1	2.1	0.6	1.6	1.5	2.6	-43.4	12.2	-12.8				
	Rmse	4.5	3.9	5.2	2.2	246.7	80.7	176.5	94.8	2.5	7.0	3.2	3.8	0.9	3.8	1.3	1.8	45.0	36.5	55.9	23.0	Rmse	14.3	17.5	8.8
	Rmse-%	18.0	155.9	28.7	49.8	41.8	193.7	42.5	71.0	8.8	22.4	11.2	15.7	4.0	16.6	6.0	8.3	17.0	140.9	28.8	52.7				
Hilatieto ja Motti-malli	harha	0.8	0.3	0.5	0.3	44.3	4.9	24.2	15.9	0.5	1.0	0.6	0.8	0.3	0.6	0.3	0.7	10.4	9.6	8.2	5.9	harha	3.5	2.0	2.5
	harha-%	3.1	6.0	3.0	6.3	6.2	5.0	5.5	9.1	1.8	3.2	2.2	3.5	1.5	2.7	1.3	3.2	4.0	25.9	4.8	12.0				
	Rmse	1.0	0.4	0.7	0.4	58.4	7.2	31.1	24.5	0.8	1.5	0.9	1.1	0.4	0.8	0.4	0.8	13.3	11.7	10.6	7.6	Rmse	4.2	2.7	3.3
	Rmse-%	3.9	8.1	4.1	8.3	8.1	7.3	7.0	14.0	3.1	4.9	3.5	4.6	2.0	3.3	1.9	3.8	5.2	31.5	6.2	15.4				

Taulukko 3. Puustotunnusten virheet ja harhat reunahilojen suodatuksen jälkeen. Hilatieto on hieman harhaisempaa, mutta virhe on vastaavasti pienentynyt monissa keskitunnuksissa. Runkoluvun osalta parannus on merkittävin. *Laskennassa runkojen tilavuuksia ei ole määritetty runkokäyrältä vaan tässä tapauksessa laskettu Laasasenahon (1982) kahden parametrin tilavuusyhtälöllä, n=150.

		G, m2/ha				N, kpl/ha				DG, cm				HG, m				V*, m3/ha				Puulajiosuudet, %			
		KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	KAIKKI	MÄNNYT	KUSET	LEHTIP.	MÄNNYT	KUSET	LEHTIP.	
Hakkuukonetieto ja Motti-malli	harha	3.7	-1.3	4.4	0.7	38.6	-35.6	65.8	8.5	1.1	0.3	1.5	0.6	0.5	0.5	0.3	0.6	-4.8	-22.2	18.2	-0.8	harha	-8.8	8.2	0.6
	harha-%	15.0	-52.8	24.3	15.4	6.5	-85.5	15.8	6.4	3.8	0.8	5.4	2.6	2.3	2.0	1.3	2.7	-1.8	-85.8	9.4	-1.7				
	Rmse	5.8	3.5	6.4	2.4	177.3	64.8	167.3	72.7	2.6	6.5	3.4	3.9	1.0	3.4	1.4	2.0	44.2	42.2	53.2	21.7	Rmse	16.4	17.9	8.1
	Rmse-%	23.3	140.6	35.4	54.5	30.0	155.6	40.2	54.5	9.1	21.0	11.7	16.0	4.5	14.9	6.3	9.0	16.8	162.9	27.4	49.6				
Hakkuukonetieto ja hilatiedon keskitunnukset	harha	-0.9	-2.2	1.3	0.0	-130.4	-58.0	-38.0	-34.8	0.7	-0.4	1.3	0.5	0.2	-0.1	0.2	0.1	-9.1	-13.3	11.3	-7.2	harha	-5.0	6.5	-1.5
	harha-%	-3.6	-85.8	7.1	-0.4	-22.1	-139.3	-9.1	-26.0	2.6	-1.1	4.4	1.9	1.1	-0.6	0.8	0.3	-3.5	-51.3	5.8	-16.4				
	Rmse	4.4	4.1	4.8	2.0	254.1	81.6	180.7	91.4	2.4	6.9	3.1	3.8	0.8	3.6	1.3	1.9	43.8	37.5	50.6	21.9	Rmse	14.4	17.1	8.4
	Rmse-%	17.6	162.3	26.4	46.3	43.0	195.8	43.5	68.5	8.2	22.1	10.9	15.7	3.7	16.0	6.0	8.4	16.6	144.8	26.0	50.1				
Hilatieto ja Motti-malli	harha	0.8	0.3	0.6	0.3	46.8	5.3	27.2	14.9	0.6	1.3	0.6	0.9	0.3	0.6	0.3	0.7	9.9	9.5	10.7	7.1	harha	3.8	2.5	2.4
	harha-%	3.1	1.8	12.9	1.0	6.5	1.2	16.2	66.2	1.8	4.6	2.5	0.3	1.5	2.8	1.3	14.8	25.2	5.2	21.0	7.1				
	Rmse	1.1	0.4	0.8	0.4	62.5	7.9	35.8	25.4	0.9	1.9	0.9	1.3	0.4	0.8	0.4	0.9	14.2	11.9	14.0	9.2	Rmse	4.8	3.3	3.2
	Rmse-%	4.1	2.4	18.5	1.3	8.7	1.7	21.3	112.9	2.7	6.9	3.6	0.5	1.9	3.6	1.8	18.4	36.2	6.5	27.6	9.2				

4.3.2 Runkoluku

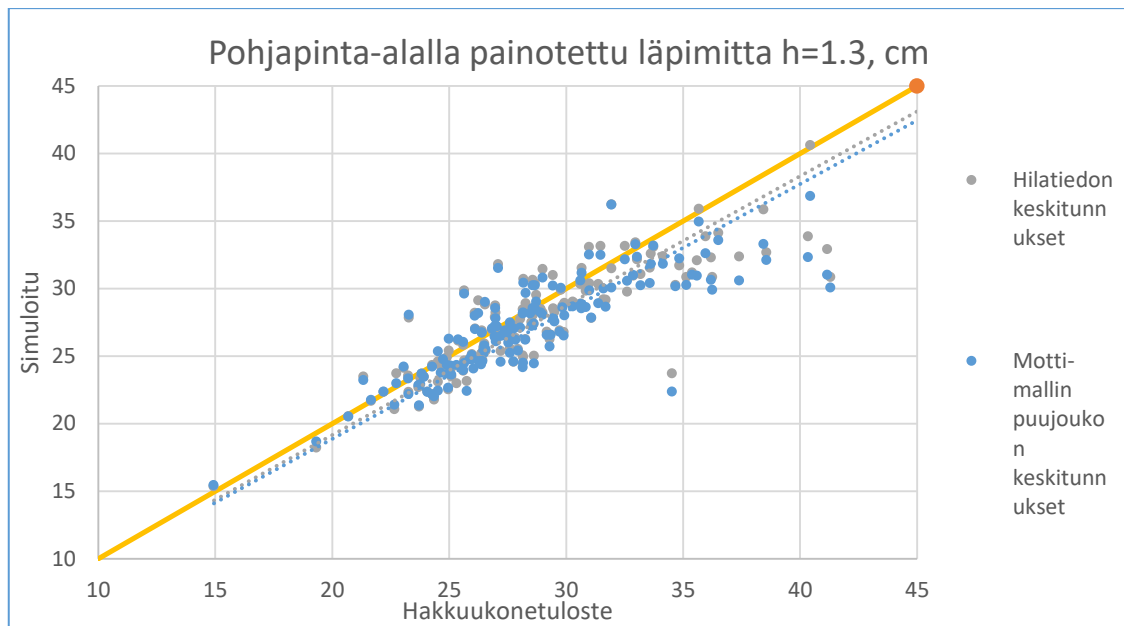
Runkoluvun tarkastelussa ilmenee hilatiedon epätarkkuus verrattuna hakkuukonetietoon. Motti-malli aliarvioi hilatiedon runkolukua systemaattisesti (Kuva 8); harha hilatiedon ja mallin välillä oli 3 % ja RMSE 4 % (Taulukko 2). Tässä aineistossa runkoluvun aliarvio pienensi virhettä hakkuukonetietoon verrattaessa, sillä harha mallin ja hakkuukonetiedon välillä oli -14% sekä RMSE 34 %, kun hilatiedon ja hakkuukonetiedon harha oli -21 % sekä RMSE 42 %. Runkolukua voidaan virheen perusteella pitää kokonaistilavuuden jälkeen heikoiten estimoituna puustotunnukseksi tässä aineistossa.



Kuva 8. Runkoluvun suhde todellisten keskitunnusten ja estimointien välillä. Motti-mallin keskitunnukset ovat lähempänä todellisia kuin hilatiedon keskitunnukset.

4.3.3 Läpimitta

Pohjapinta-alalla painotetun keskiläpimitan estimoinnissa Motti-malli tuotti hieman epätarkempia tuloksia, kuin alkuperäiset hilatiedon keskitunnukset (Taulukko 2). Hilan ja mallin välinen harha oli 0,5 cm (1,8 %) sekä virhe 0,8 cm (3,1 %), ja mallin ja hakkuukonetiedon välinen harha 1,4 cm (4,9 %) ja virhe 3,0 cm (10,3 %). Männyllä tulos oli tarkin; harha 0,5 cm (1,6 %) ja kuusella heikoin 1,8 cm (6,4 %). Läpimitan ennustaminen onnistui paremmin kuvioilla, jossa läpimitta oli pieni, mutta läpimitan kasvaessa sen ennustus heikkeni (Kuva 9).



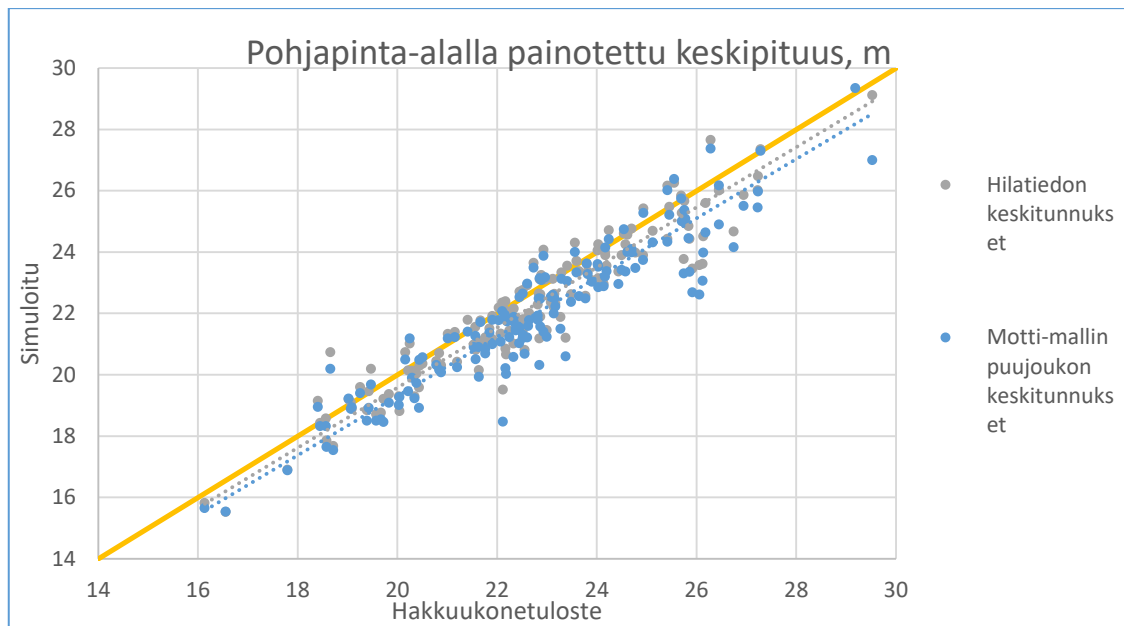
Kuva 9. Pohjapinta-alalla painotetun läpimitan suhde. Estimoidut arvot aliarvioivat keskimäärin läpimittoja.

4.3.4 Pituus

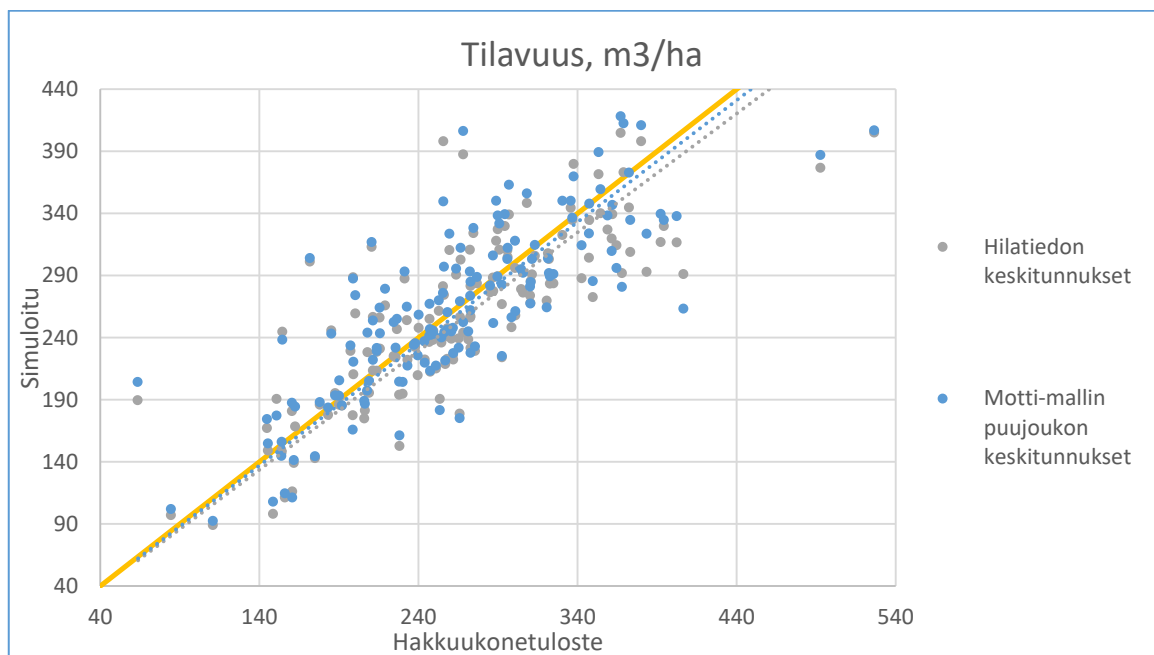
Pituuden estimoinnissa mallin ja hakkuukonetiedon välillä oli lievää aliarviota (Taulukko 2). Harha oli 0,8 m (3,4 %) ja RMSE 1,2 (5,2 %). Hilatieiden ja hakkuukonetiedon välinen pituuden harha oli 0,5 m sekä RMSE 0,9 m (4,0 %). Malli ennusti pituuden tunnuksia keskimäärin heikommin kuin hila-tieto (Kuva 10). Kun verrattiin hilan ja mallin puulajikohtaisia keskitunnuksia hakkuukonetietoon, mallin harhat olivat männyn ja lehtipuiden osalta suuremmat, mutta virhe sitä vastoin pienempi kuin hila-tiedossa. Kuusen kohdalla hila-tiedon virhe oli pienempi ja harha suurempi kuin mallin keskitunnuksissa.

4.3.5 Tilavuus

Puuston tilavuuteen on kaikilla edellä mainituilla puustotunnuksilla vaikutusta, mutta selkeimmin tilavuus korreloi pohjapinta-alan kanssa. Mallin ja hakkuukonetiedon välinen harha oli 0,4 m³/ha (0,1 %) sekä RMSE 45,2 m³/ha (17,1 %). välillä harha on suurinta männyllä, jossa malli yliarvioi puulajin osuutta 20 m³/ha (80 %). Tilavuuden estimoinnissa esiintyi paljon hajontaa (Kuva 11).



Kuva 10. Pohjapinta-alalla painotetun pituuden todellisten arvojen ja estimointien suhde. Motti-malli tuotti keskimäärin harhattomampia läpimittoja kuin hilatieto, vaikka molemmat aliarvioivat todellisia tunnuksia.



Kuva 11. Kuviodien kokonaistilavuuksien todellisten arvojen ja estimointien suhde.

4.3.6 Puulajiosuudet

Puulajiosuudet muuttuivat Motti-mallin simulaation jälkeen. Männyn harha kasvoi 3 % ja RMSE 4 % verratessa mallin tulosta hilatietoon. Hakkuukoneaineiston ja mallin väliset virheet olivat huomattavia.

4.3.7 Keskitunnukset

Taulukko 4. Kuvioiden keskitunnukset kolmesta eri lähteestä päätehakkuukuvioiden, jossa G on pohjapinta-ala, N runkoluku, DG pohjapinta-alalla painotettu keskiläpimitta, HG pohjapinta-alalla painotettu pituus sekä V tilavuus hehtaarilla. Toteutuneet arvot laskettu hakkuukonetulosteesta, Motti-mallilla luotuna hilatiedosta sekä hilatiedon alkuperäiset keskitunnukset.

	G, m ² /ha	N, kpl/ha	DG, cm	HG, m	V, m ³ /ha	Osuus
<i>Toteutunut</i>	25.0	591	28.7	23	264	100 %
<i>Motti-malli</i>	25.1	672	27.3	22	263	100 %
<i>Hilatieto</i>	24.6	716	27.7	22	257	100 %
Männnyt						
<i>Toteutunut</i>	2.5	42	31.3	23	26	10.5 %
<i>Motti-malli</i>	4.6	94	30.6	22	47	18.9 %
<i>Hilatieto</i>	4.5	99	31.2	23	37	15.5 %
Kuuset						
<i>Toteutunut</i>	18.1	416	28.7	22	194	71.8 %
<i>Motti-malli</i>	15.9	419	26.9	22	172	63.6 %
<i>Hilatieto</i>	15.9	443	27.2	22	171	64.7 %
Lehtipuut						
<i>Toteutunut</i>	4.3	134	24.4	22	44	17.7 %
<i>Motti-malli</i>	4.5	159	23.3	21	45	17.6 %
<i>Hilatieto</i>	4.3	175	23.6	22	49	19.7 %

Puujoukkojen keskitunnukset ovat runkolukua lukuun ottamatta melko yhteneväiset (Taulukko 4). Merkittäviä eroja löytyy hehtaariohjeista runkoluvuista. Myös puulajiosuuksissa on huomattavia poikkeuksia; pohjapinta-aloissa männyllä havaittavissa yliarviota ja sitä vastoin kuusella aliarviota. Männyn tilavuuden osuus on Motti-mallissa 3,4 prosenttiyksikköä suurempi hilatiedon keskiarvoihin verrattuna. Ero todelliseen on vieläkin suurempi, 8,4 prosenttiyksikköä.

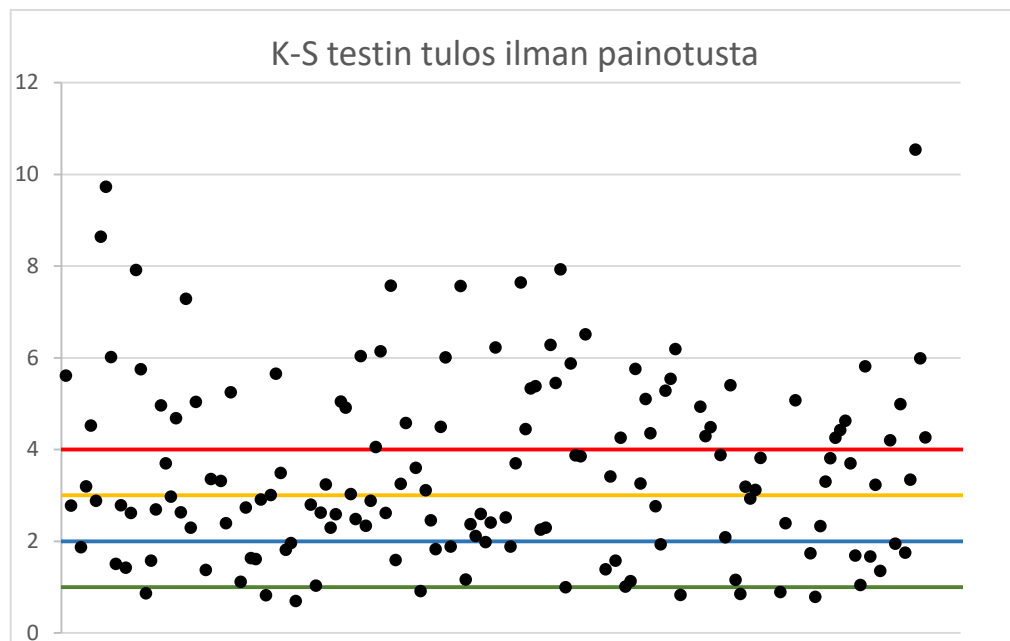
4.4 Runkolukusarjojen mallinnus

4.4.1 Yhteensopivuustestit

Valitut yhteensopivuustestit olivat Kolmogorov-Smirnov sekä Reynolds'n virheindeksi (Reynolds ym. 1988). Kolmogorov-Smirnov tulkitsee kumulatiivisten suhteellisten runkolukukäyrien suurinta luokkien välistä etäisyyttä D. Lisäksi etäisyydelle voidaan laskea kriittinen arvo, joka kuvastaa arvon D tilastollista merkitsevyyttä. Kriittisen arvon

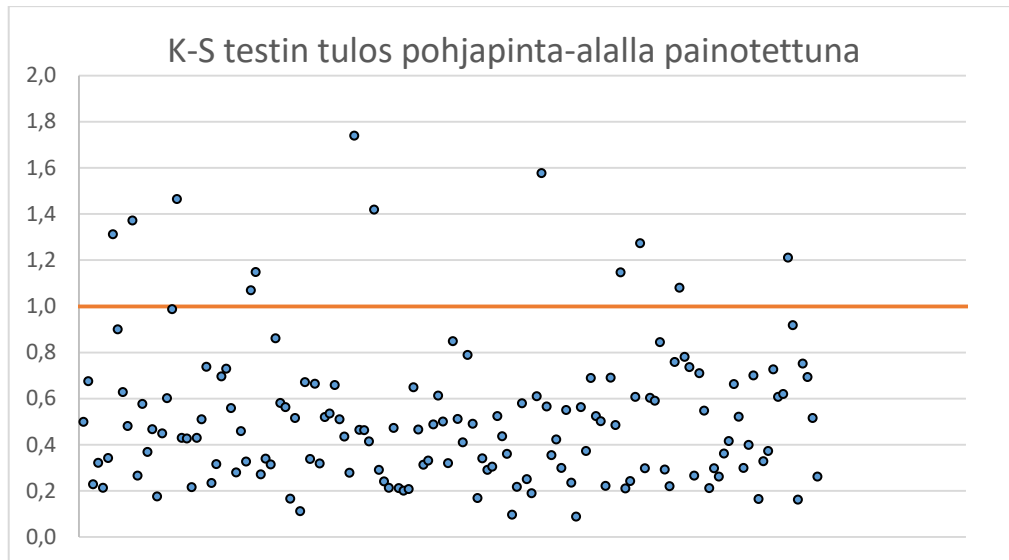
laskukaavan yksi muuttujista oli runkoluku. Vaikka Reynolds'n virheindeksi ei ole virallinen tilastollinen menetelmä, se on kehitetty juuri runkolukusarjojen vertailua varten sekä se tarjoaa mahdollisuuden läpimittaluokkien painottamiseen. Lisäksi sitä on käytetty useassa aiemmassa tutkimuksessa runkolukusarjojen vertailussa.

Kolmogorov-Smirnov testin osamäärä oli yhdeksällä kuviolla alle yksi, kun testin merkitsevyystaso oli 0,05. Tämä tarkoittaa yhdeksää prosenttia koko aineistosta. Tham (1988) esitti, että K-S testiä tulisi tulkita siten, ettei osamäärältään yli neljän kuvioita tulisi hyväksyä. Tällä menettelyllä testin läpäisi 96 kuviota (64 %) (Kuva 12).



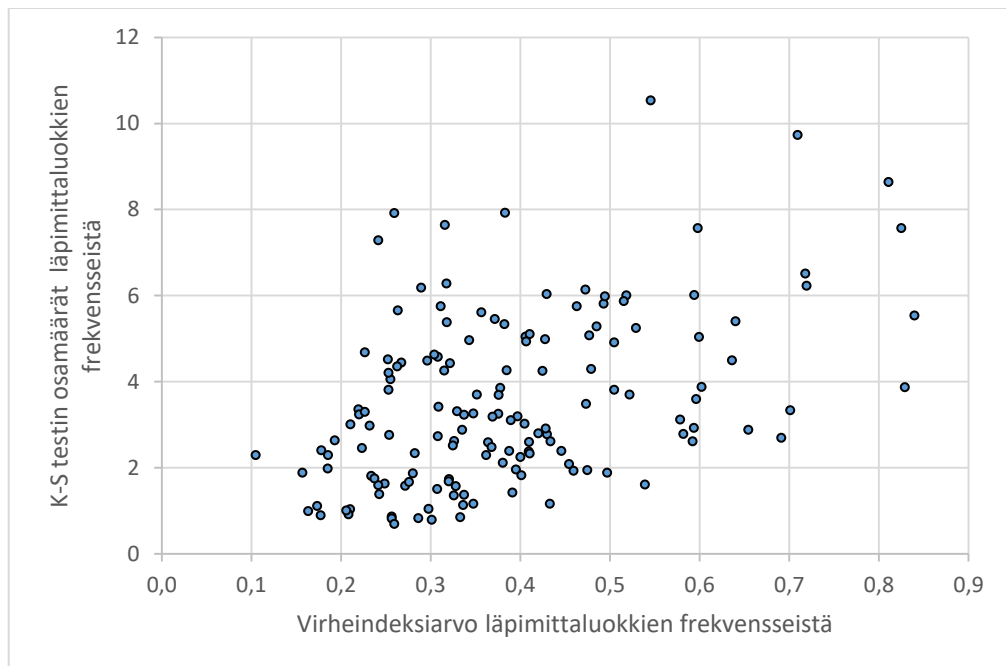
Kuva 12. Kolmogorov-Smirnov -testin osamäärät kuvioittain. Päätehakkuukuviot esitetty harmaana havaintoina, vihreä jana kuvaa osamäärän rajaa yksi, sininen rajaa kaksi, keltainen rajaa kolme ja punainen rajaa neljä.

Kun runkoluvun tilalle vaihdettiin pohjapinta-ala, voitiin testin painotusta kääntää enemmän suuriläpimittaisiin puihin. Yhteensä 83 % kuvioista läpäisi testin alle yhden osamäärällä (Kuva 13).



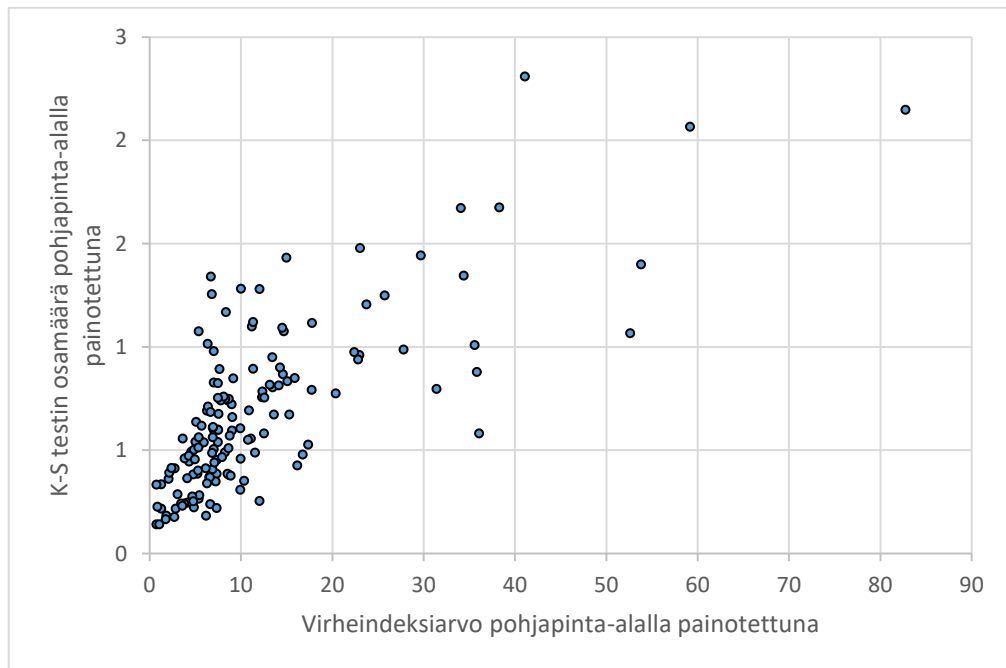
Kuva 13. K-S testin tulokset kuvioittain, kun testissä käytettiin pohjapinta-alalla painotettuja arvoja. Kuvioista 124 kappaletta läpäisi testin alle yhden osamäärällä, joka ilmentää 95 % merkitsevyystasoa.

K-S testin ja virheindeksin keskinäisen korrelaation tarkastelulla haluttiin validoida testien toimivuus. Testien laskentamenetelmä on erilainen ja siten niillä tutkitaan eri kriteerejä; K-S testi etsii luokkien väliltä suurinta lukumääräeroa, kun virheindeksi summaa kaikki erotukset yhteen. Testien arvojen korrelaatioita vertailemalla voidaan todeta, että vaikka virheindeksin antama arvo on kuviolle hyvä, K-S testin osamäärät vaihtelivat suuresti. Testien keskinäinen korrelaatio oli 0,09 (Kuva 14).



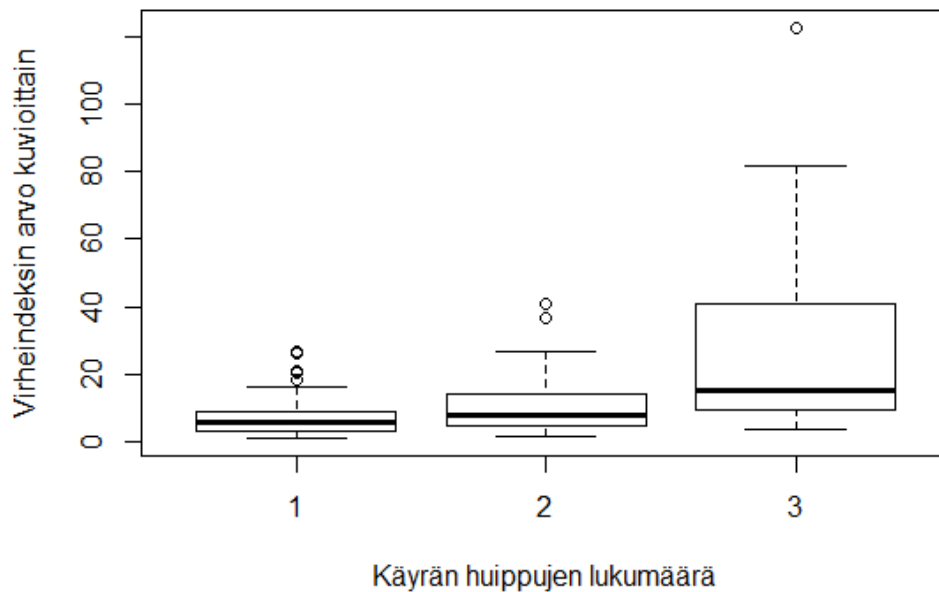
Kuva 14. Kolmogorov-Smirnov-testin ja virheindeksin tulosten keskinäinen suhde. Testien tulosten korrelaatiota on vaikea hahmottaa (0,09).

Testien keskinäinen korrelaatio parani huomattavasti, kun läpimittaluokkien frekvenssien tilalla käytettiin frekvenssin edustamaa pohjapinta-alaa (Kuva 15). Painotus siirtyi silloin suurempiin puihin ja testien välinen korrelaatioksi saatiin 0,71.



Kuva 15. Kolmogorov-Smirnov -testin ja virheindeksin tulosten keskinäinen suhde, kun molemmissa testeissä käytettiin pohjapinta-alalla painotettuja arvoja. Korrelaatio 0,71.

Jakaumien tarkkuutta arvioitiin arvioimalla subjektiivisesti huippujen lukumäärä aidon aineiston runkolukusarjalla. Tällä haluttiin tutkia sitä, onnistuuko jakauman estimointi paremmin homogeenisessä metsikössä. Tuloksia vertailtiin pohjapinta-alalla painotettuun virheindeksiin. Neliö- ja janakaaviosta voi havaita, kuinka yhden ja kahden huipun kuvioilla virheindeksi on pieni verrattuna kolme- tai useampihuippuisiin jakaumiin (Kuva 16). Mallin, jossa virheindeksi toimii huippujen selittäjänä, korrelaatio oli 0,11. Virheindeksi, josta alle 15 cm läpimittaiset puut olivat suodatettu pois, korrelaatio oli 0,47. Käyrien lukumäärää verrattiin myös katkontatulosten kokonaistilavuuden virheisiin sekä pölkkyjen lukumäärien virheisiin. Virheinä käytettiin absoluuttisia arvoja. Mallin korrelaatio oli -0,09. Koska virheindeksimenetelmässä lukumäärät sijoitetaan kumulatiiviselle käyrälle, jossa runkojen lukumäärät ilmoitetaan osuuksina yhdestä, oli näin ollen selkeämpää verrata hehtaariohtaisia tilavuus- ja pölkky-määriä virheindeksin arvoon.



Kuva 16. Ruutu- ja janakaavio runkolukusarjojen huippujen lukumääristä suhteessa virheindeksin arvoon. Virheindeksistä on suodatettu alle 15 cm läpimittaiset puut. Korrelaatio oli 0,48.

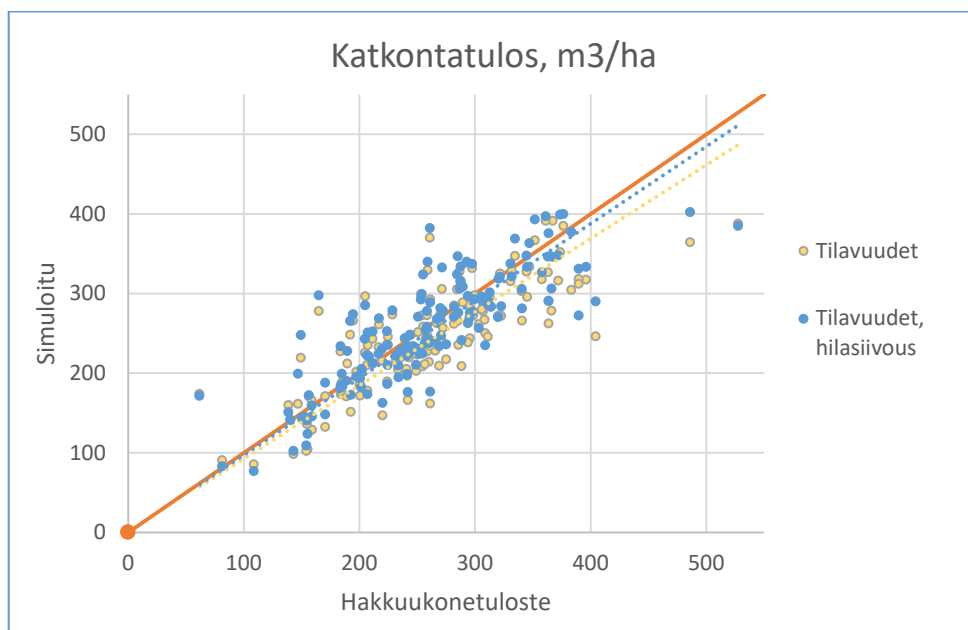
4.5 Katkonnan estimoinnin tarkkuus päätehakkuukuvioilla

Katkonnan tuloksia vertailtiin kahden eri toteutustavan mukaan riippuen laskentaan käytettyjen hilojen määrästä. Alkuperäisessä laskennassa käytettiin kaikkia hiloja (Taulukko 5) ja vertailussa leikkautuneet, hilan kokonaispinta-alasta alle 75 % kokoiset hilat, suodatettiin pois laskennasta (Taulukko 6). Kokonaistilavuuden tarkastelussa menestyi paremmin malli, jossa reunahilat olivat suodatettu pois kuvioden reunoilta sekä pölkyt ja tilavuudet olivat kerrottu tilavuuksien erosta lasketun skaalauskerroimen mukaan. Mielenkiintoisimmat tunnuksat olivat pölkkyjen tilavuuksien ja keskijäreysien harhat sekä virheet, mitkä ilmentävät puuston samankaltaisuutta.

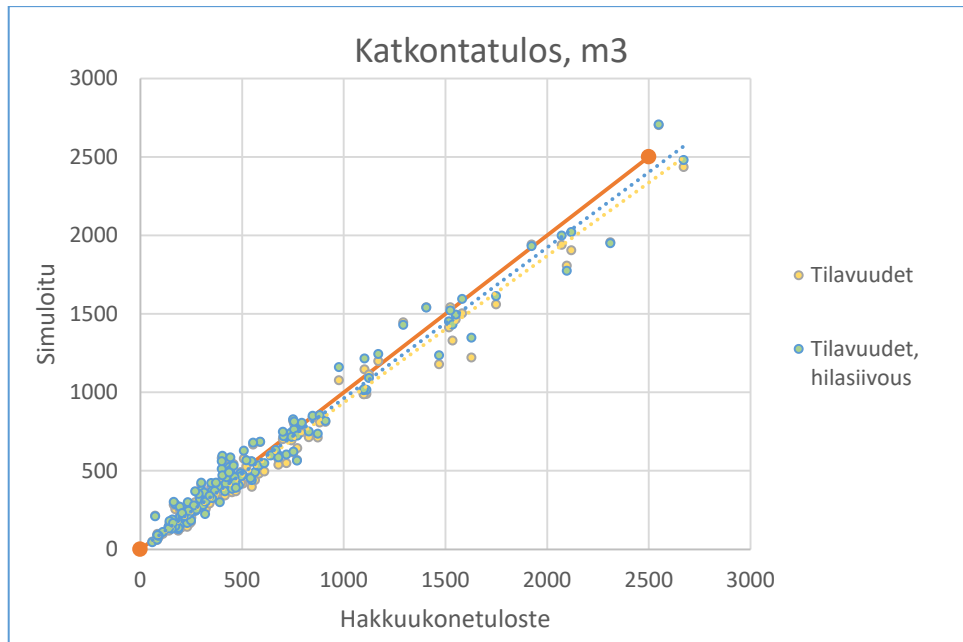
Alkuperäisen laskennan kokonaistilavuuden RMSE oli päätehakkuukuvioilla 46,2 m³/ha (17,8 %) ja harha 15,5 m³/ha (6,0 %) (Taulukko 5). Puulajikohtaisesti parhaiten menestyi kuusi ja heikoiten mänty. Männyn kokonaistilavuuden RMSE oli 40,4 m³/ha (154,9 %) ja harha -18,0 m³/ha (-69,1 %). Kuusen kokonaistilavuuden RMSE oli puolestaan 58,8 m³/ha (30,9 %) ja harha 32,1 m³/ha (16,9 %). Lehtipuiden kokonaistilavuuden RMSE oli 22,7 m³/ha (52,3 %) ja harha vertailun pienin 1,4 m³/ha (3,3 %). Malli yliarvioi systemaattisesti männyn ja aliarvioi kuusen osuutta kuvioilla. Männyn pienen todellisen puulajiosuuden vuoksi suhteelliset harhan ja keskivirheen arvot ovat poikkeuksellisen suuria. Kuusitukkien harha oli aineistossa myös suuri, 31,5 m³/ha (20,0 %). Suurta aliarviota ei selitä pelkästään mäntytukkien määrä, sillä niiden harha oli vain -

13,3 m³/ha. Runkolukusarjoja tarkastellessa selvisi, että usealle kuviolle mäntyä oli estimoitu liikaa. Jopa niille kuvioille, joilla mäntyä ei kasvanut, oli estimoitu suuriakin runkoja. Puuston tukkiprosenttien virhe oli 3,4 % ja RMSE 6,2 %. Puulajikohtaisesti tukkiprosenttien harhat ja keskivirheet olivat männyllä 3,8 % ja 9,9 %, kuusella 2,9 % ja 7,9 % sekä koivulla 3,7 % ja 12,5 %. Simuloitu tieto aliarvioi puuston keskijäreyttä (Taulukko 5), sillä tukkien keskijäreyden harha oli 15,0 litraa (6,5 %) ja RMSE 34,5 litraa (34,5 %).

Reunahilojen suodatuksen jälkeen kokonaistilavuuden RMSE oli 42,3 m³/ha, (16,3 %) ja harha 3,8 m³/ha (1,5 %) (Taulukko 6). Kokonaistilavuuden RMSE parani siten 1,5 prosenttiyksikköä, mutta ilmeisempi parannus tapahtui harhan suhteen, 4,5 prosenttiyksikköä. Männyn kokonaistilavuuden keskivirhe oli 40,3 m³/ha (154,5 %) ja harha -18,9 m³/ha (-72,5 %). Kuusen tilavuuden keskivirhe oli 53,3 m³/ha (28,0 %) sekä harha 21,8 m³/ha (11,5 %). Lehtipuiden kokonaistilavuuden RMSE oli 21,7 m³/ha (50,1 %) ja harha 0,9 m³/ha (2,1 %). Mäntytukkien osalta harhat sekä virheet kasvoivat hieman, mutta kusitukkien harha väheni arvoon 22,2 m³/ha (14,1 %). Kuusikuitupölkkyjen osalta harha oli -0,3 m³/ha (-1,0 %), mikä osoittaa pieniläpimittaisten kuusirunkojen yliarvion kuvioilla. Puuston tukkiosuuksien keskivirhe oli 5,4 % ja harha 2,4 %. Tukkiprosenttien harhat sekä keskivirheet olivat männyllä, kuusella ja lehtipuilla 3,8 % ja 9,9 %, 2,9 % ja 7,9 % sekä 3,7 % ja 12,5 %.

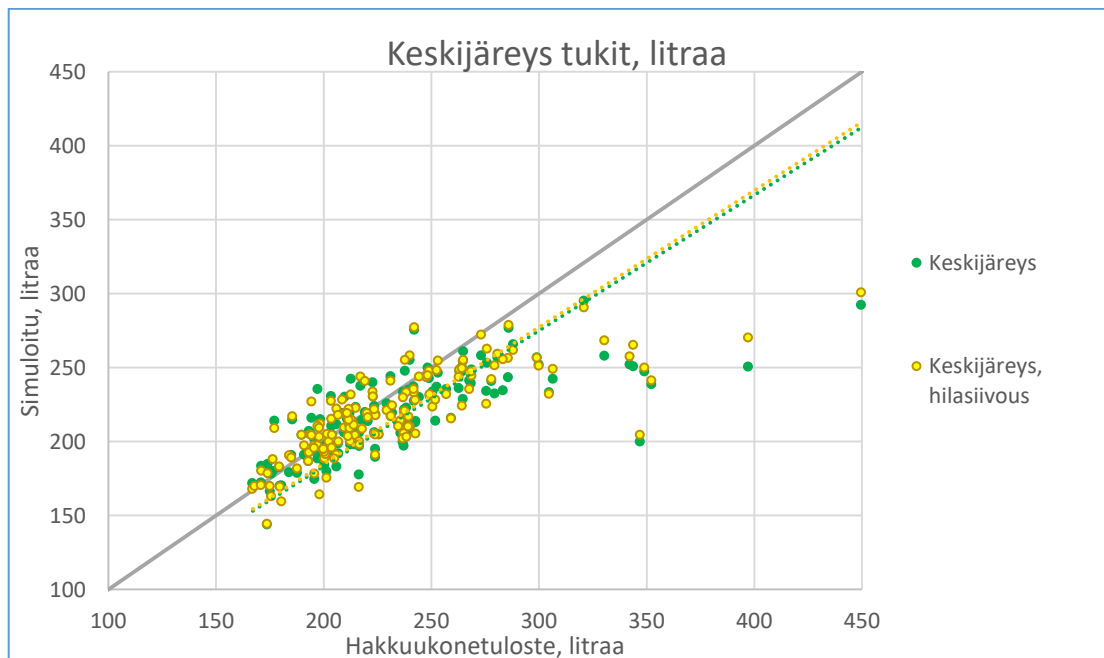


Kuva 17. Hehtaarikohtaisten tilavuuksien toteutuneiden ja simuloitujen katkontatulosten suhde. Katkontatulos oli reunahilojen suodatuksen jälkeen harhattomampi.



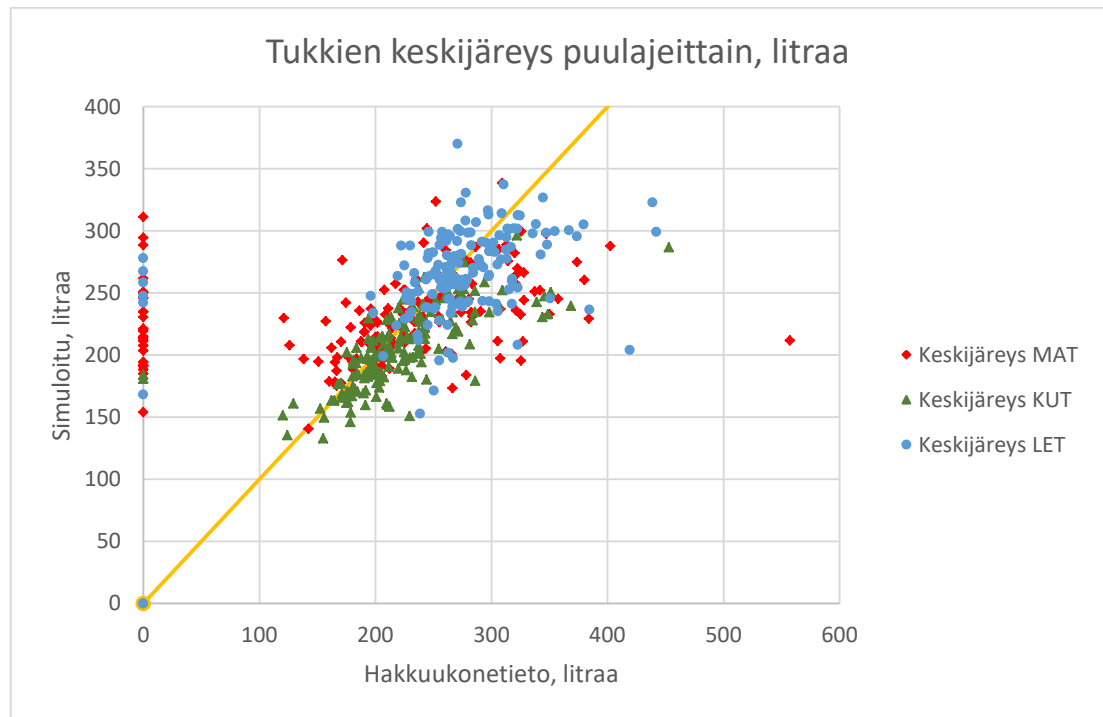
Kuva 18. Kokonaistilavuuden toteutuneiden ja simuloitujen katkontatulosten suhde. Tilavuuden ennustaminen on onnistunut hyvin kokonaistilavuudesta riippumatta, reunahilojen suodatuksen jälkeen hieman paremmin.

Katkontatuloksen tarkkuudessa on hieman eroa eri menetelmien välillä (Kuva 17). Hehtaarikohtaisissa tilavuustunnuksissa hajonnan voi erottaa selkeästi, mutta absoluuttisissa tuloksissa (Kuva 18) ilmenee, kuinka suuret kuvioiden kokonaistilavuudet eivät ole yhtä virheellisiä. Kuviokoon suuretsa myös hilojen määrä kuviolla kasvaa, millä on merkitystä tilavuuksien tarkkuuden laskennassa.



Kuva 19. Toteutuneen ja simuloitujen keskijäreysien suhde. Simulaatio aliarvioi keskijäreysiä etenkin järeämmässä puustossa.

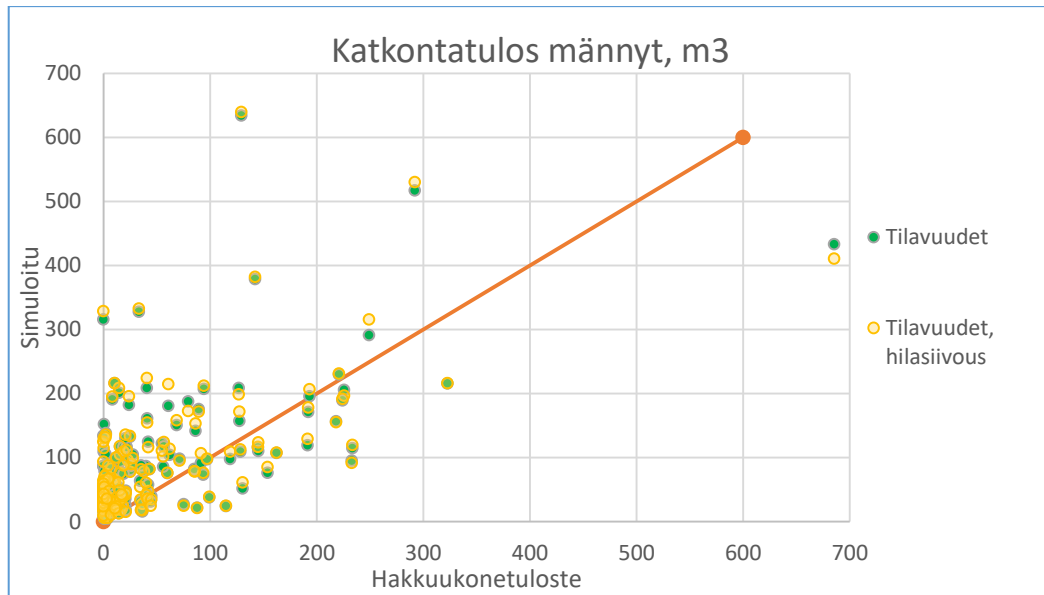
Tukkipuiden kuviokohtainen keskijäreys oli tarkinta silloin, kun keskijäreys oli alle 250 litraa (Kuva 19). Kuvioilla, joilla keskijäreys oli suurempaa, ei simulointi kyennyt ennustamaan tulosta yhtä tarkasti.



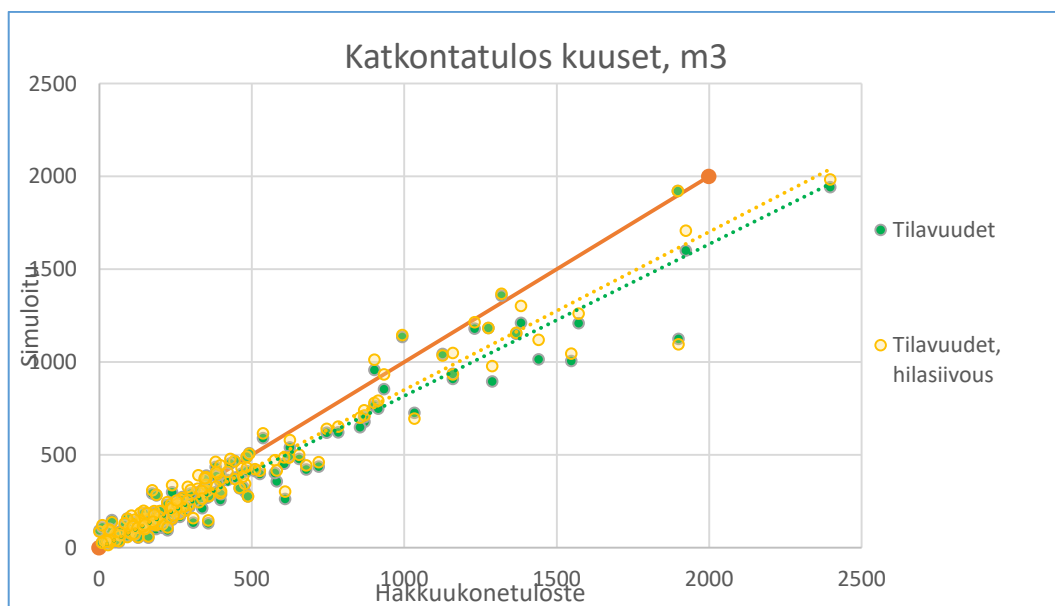
Kuva 20. Tukkien keskijäreys puulajeittain jokaiselta kuviolta. Kuusten ja lehtipuiden keskijäreiden estimointi on ollut suhteellisen harhatonta. Männyllä esiintyi eniten hajontaa.

Puulajeittain tukkien keskijäreiden estimointi onnistui heikoiten männyllä. Hajontaa esiintyi silti runsaasti jokaisella puulajilla (Kuva 20).

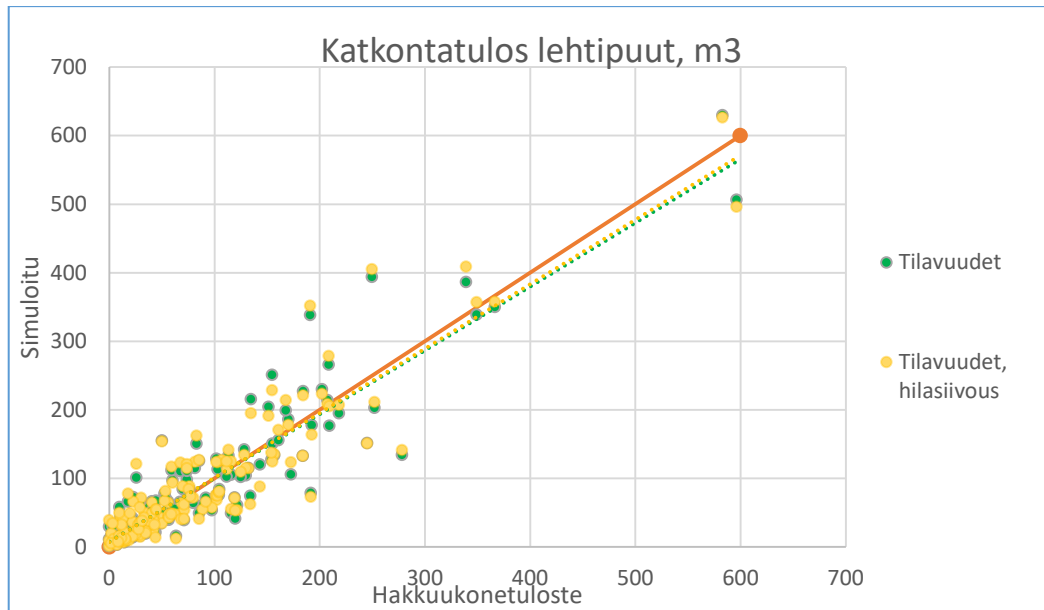
Katkonnan tuloksissa heikoiten estimoitu puulaji oli molemmilla menetelmillä mänty (Kuva 21). Männyä tilavuutta yliarvioitiin useimmilla kuvioilla ja vastaavasti kuusen tilavuutta aliarvioitiin (Kuva 22). Lehtipuiden estimaatti sisälsi paljon hajontaa, mutta oli suhteellisen harhatonta (Kuva 23). Männyä osuus oli suuri niilläkin kuvioilla, missä sitä ei hakkuukonetulosten mukaan ollut kaadettukaan. Tarkemmassa tarkastelussa ei löydetty yhteyttä männyä osuuden yliarvioon suhteessa kokonaistilavuuteen tai pölkkyjen lukumäärään.



Kuva 21. Männyn katkonnan tilavuuksien suhde simuloitun ja toteutuneen välillä. Mäntypuutavarylajeja on katkottu simulaatiossa runsaasti, vaikkei niitä todellisuudessa ollut havaittu.



Kuva 22. Kuusen katkonnan tilavuuksien suhde simuloitun ja toteutuneen välillä. Kuusipuutavarylajien tilavuuksissa oli paljon aliarviota. Reunahilojen suodatus paransi hieman tulosta.



Kuva 23. Lehtipuiden katkonnan tilavuuksien suhde simuloitun ja toteutuneen välillä. Tilavuudet olivat suhteellisen harhattomia, mutta kuviokohtaisten estimaattien välillä esiintyi runsaasti hajontaa.

Taulukko 5. Katkonnan tilavuuksien ja pölkkyjen lukumäärien harhat ja keksineliöjuuren keskivirheet. Tilavuudet hehtaariohtaisina kuutiometreinä (m³/ha) paitsi keskijäreuden tilavuudet litroina (l). MAT= Mäntytukki, MAK = mäntykuitu, KUT= Kuusitukki, KUK = Kuusikuitu, LET = Lehtipuutukki, LEK = Lehtipuukuitu

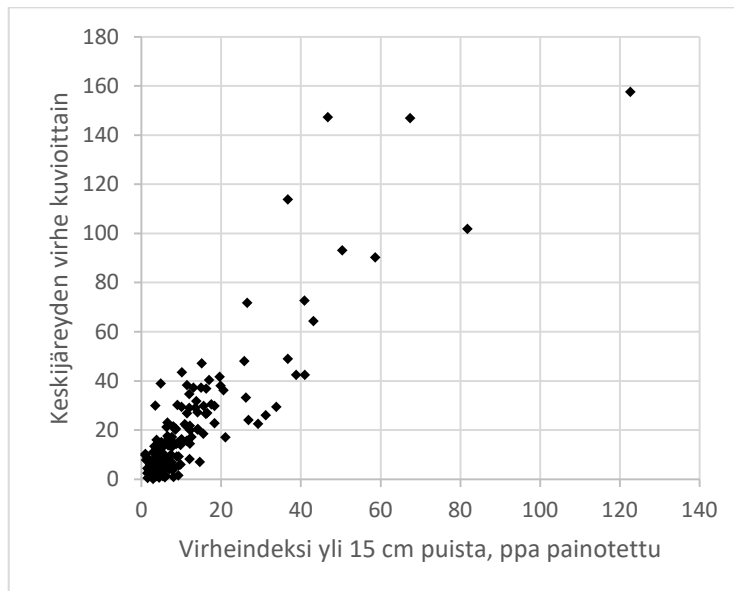
Päätihakkuukuviot, n = 150																									
	Tilavuus MAT	Pölkylt MAT	Tilavuus MAK	Pölkylt MAK	Kokonaistilavuus MA	Pölkylt yhteensä MA	Tilavuus KUT	Pölkylt KUT	Tilavuus KUK	Pölkylt KUK	Kokonaistilavuus KU	Pölkylt yhteensä KU	Tilavuus LET	Pölkylt LET	Tilavuus LEK	Pölkylt LEK	Kokonaistilavuus LE	Pölkylt yhteensä LE	Kokonaistilavuus	Pölkylt yhteensä	Keskijäreys tukit, litraa	Keskijäreys kuidut, litraa	Keskijäreys MAT, litraa	Keskijäreys KUT, litraa	Keskijäreys LET, litraa
bias	-13.3	-60.4	-4.7	-114.3	-18.0	-174.8	31.5	102.0	0.6	7.5	32.1	109.5	2.3	5.7	-0.8	-46.3	1.4	-40.6	15.5	-105.9	15.0	2.5	12.7	18.6	11.9
bias-%	-58.7	-64.1	-137.9	-144.9	-69.1	-100.9	20.0	14.8	1.7	1.3	16.9	8.7	15.3	11.1	-2.9	-7.9	3.3	-6.3	6.0	-5.1	6.5	4.8	5.2	8.3	4.3
RMSE	35.2	141.8	6.7	159.6	40.4	284.9	52.2	226.9	12.3	230.1	58.8	408.4	11.9	34.3	14.2	296.8	22.7	312.6	46.2	465.9	34.5	5.8	58.1	36.7	43.6
RMSE-%	154.8	150.4	196.0	202.2	154.9	164.5	33.1	32.9	37.5	40.3	30.9	32.4	80.1	67.3	49.9	50.4	52.3	48.9	17.8	22.5	14.9	10.9	23.7	16.3	15.7

Taulukko 6. Katkonnan tilavuuksien ja pölkkyjen lukumäärien harhat ja keksineliöjuuren keskivirheet, kun reunahilojen suodatus ja kuvioiden skaalaus olivat käytössä.

Päätehakkuukuviot, n = 150																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
	Pölkyl yhteensä MA						Pölkyl yhteensä KU						Pölkyl yhteensä LE						Pölkyl yhteensä		Keskijäreys, tukit, litraa			Keskijäreys MAT, litraa			Keskijäreys KUT, litraa		Keskijäreys LET, litraa																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																						
	Tilavuus MAT	Pölkyl MAT	Tilavuus MAK	Pölkyl MAK	Kokonaistilavuus MA		Tilavuus KUT	Pölkyl KUT	Tilavuus KUK	Pölkyl KUK	Kokonaistilavuus KU		Tilavuus LET	Pölkyl LET	Tilavuus LEK	Pölkyl LEK	Kokonaistilavuus LE																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		</

4.6 Puustotunnusten selitettävyys tarkkuuden suhteen

Puusto- ja kasvupaikkatunnuksia verrattiin virheindeksin arvoon sekä katkonnessa tuotettujen tukkipölkkyjen tilavuuden tarkkuuteen tutkimalla korrelaatiokertoimia (Taulukko 7). Virheindeksille laskettiin kaksi eri arvoa, joista toinen laskettiin rinnankorkeuslähimitaltaan vain yli 15 cm puille ja molemmissa käytettiin painotusta pohjapinta-alalla. Korrelaatio eri puustotunnusten välillä oli selvästi suurempaa virheindeksin arvoilla, joissa alle 15 cm puut olivat suodatettu laskennasta pois. Tukkitilavuuden virhe laskettiin katkonnan tuloksesta. Käyttämällä tukkipuiden tilavuutta katkonnan kokonaistilavuuden sijaan voitiin kohdistaa tarkkuuden laskenta kiinnostuksen kohteena oleviin puutavaralajeihin.



Kuva 24. Virheindeksin sekä tukkien keskipääteiden absoluuttisen virheen suhde. Tunnuksilla on selkeä korrelaatio, vaikka keskipääte ei sisällä informaatiota pölkkyjen tilavuuksien tai lukumäärien todellisista määräeroista.

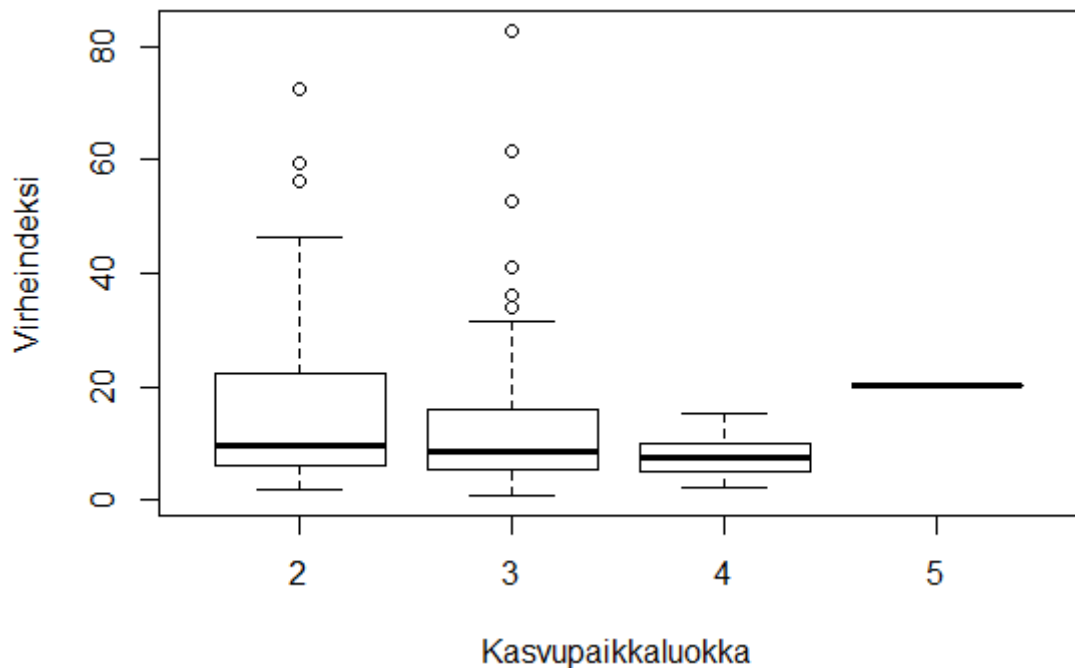
Puulajikohtaista tarkkuutta esitettiin laskemalla puulajeittain tukkien keskipääteiden tarkkuus vähentämällä suhteellinen virhe luvusta yksi, ja lukuja tukkien tilavuusmäärällä painottaen. Luvun asettuessa nollan ja yhden välille, on sen tarkoitus mukailla jakauma-astetta, vaikka laskentatapa on erilainen. Tukkien keskipääte toimii hyvin tarkkuuden määrittäjänä, sillä se ottaa huomioon sekä tilavuuden että pölkkyjen lukumäärät. Vaikka keskipääteydessä katoaa tieto molemmista, se korreloi silti hyvin yli 15 cm runkojen virheindeksin arvon kanssa (Kuva 24).

Taulukko 7 Eri selittäjien korrelaatiokertoimet eri analyysien selittäjinä. Selittäjien 95 % merkitsevyystaso on merkitty tähdellä. Virheindeksi on laskettu käyttäen kaikkia puita sekä vain läpimitaltaan yli 15 cm puita, molemmissa käyttäen pohjapinta-alapainotusta. Tukkien tilavuuden virhe on laskettu katkonnan tuloksen tukkipölkkyjen tilavuuksista. Tarkkuus on laskettu laskemalla puulajeittainen keskipäijäyden osuus tilavuuksilla painottaen. CV on variaatiokerroin.

	Virheindeksi	Virheindeksi, lpm > 15 cm	Virhe-%, tukki- tilavuus	Plaji-tarkkuus-%, tukkien keskipäijäyden
Runkoluku kpl/ha	-0.05	-0.42 *	-0.01	0.17 *
Ppa m ² /ha	0.03	0.10 *	0.16	0.39 *
Lpm, cm	0.05	0.53 *	0.18 *	0.09
Pituus, m	0.05 *	0.33 *	0.15	0.16 *
DG, cm	0.08 *	0.79 *	0.17 *	0.05
HG, m	0.09 *	0.60 *	0.19 *	0.17 *
Kuusen osuus, %	0.04	0.08 *	0.22 *	0.38 *
CV lpm	0.01	0.09 *	0.19 *	0.12
CV pituus	-0.01	0.19 *	0.17 *	0.09

Kerrointen perusteella etenkin virheindeksiin, jossa alle 15 cm puut ovat suodatettu, vaikuttaa eniten puustotunnukset pituus, pohjapinta-alalla painotettu läpimitta sekä pohjapinta-alalla painotettu pituus. Tukkitilavuuden virheprosenttiin vaikuttivat useampi muuttuja, joista suurin korrelaatio oli puuston dimensiomuuttujilla, kuusen osuudella sekä variaatiokertoimella (CV). Puulajikohtaiseen tarkkuuteen suurin vaikutus oli kuusen osuudella ja pohjapinta-alalla. Virheindeksi ei korreloinut paljoa minkään selittävän muuttujan kanssa, mutta sen sijaan suodatettu virheindeksi korreloi kaikkien muuttujien kanssa merkitsevästi. Puuston runkoluku vaikutti virheindeksin arvoon negatiivisesti, ja eniten korrelaatiota oli molemmilla virheindeksin arvoilla pohjapinta-alalla painotettujen dimensiotunnusten kanssa.

Kasvupaikkaluokalla uskottiin olevan merkitystä puuston tarkkuuden kanssa. Kasvupaikkaluokat poimittiin suoraan hilatiedon attributeista ja laskettiin kasvupaikkatyyppin keskiarvo niille kuvioille, joille ilmoitettu kasvupaikkaluokitus löytyi. Kun testattiin pohjapinta-alalla painotettuja virheindeksin arvoja ja kasvupaikkaluokkia ruutu- ja janakaaviossa, huomattiin, ettei kasvupaikkaluokka toiminut tässä tutkimuksessa hyvänä selittäjänä puuston tarkkuuden suhteen (Kuva 25). Rehevimmillä kasvupaikoilla oli runsaasti hajontaa, mutta niiden keskiarvot ovat lähellä toisiaan. Kuivan kankaan (CT) kuvioita oli tutkimuksessa mukana vain yksi kappale, joten tarkkuudesta ei voi tässä kasvupaikkaluokassa tehdä johtopäätöksiä.



Kuva 25. Virheindeksi kuvattuna ruutu-ja-jana -kuvaajassa kasvupaikkaluokittain. Kasvupaikkaluokalla ei näyttänyt olevan lineaarista yhteyttä virheindeksin suuruuteen. Kasvupaikkaluokat ovat SOLMU-luokituksen mukaisia: 2=OMT, 3=MT, 4=VT, 5=CT.

5 TULOSTEN TARKASTELUA

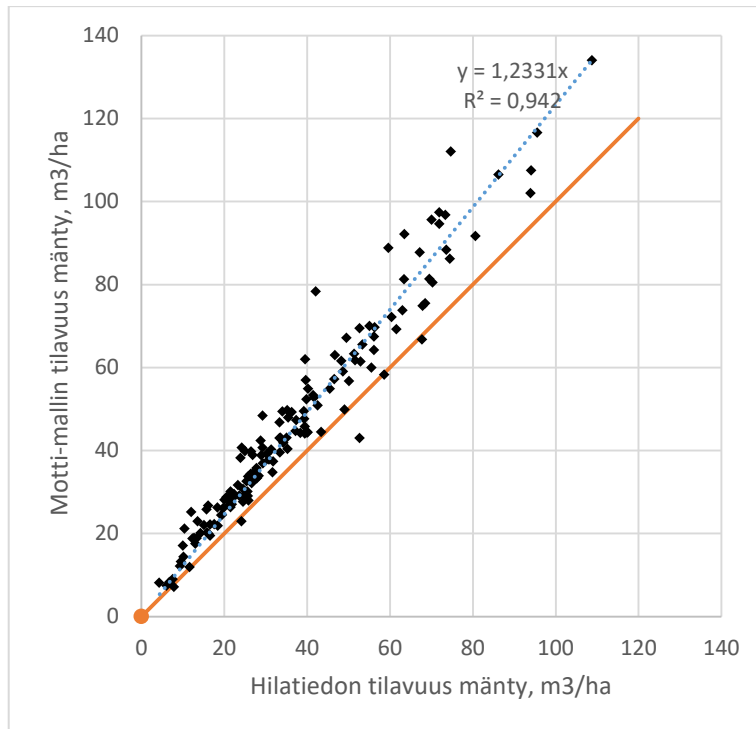
5.1 Keskitunnusten virheet

Puuston keskitunnusten vertailussa olennaista oli tunnistaa hilatiedon estimointitarkkuus sekä Motti-mallin tuottamat muutokset keskitunnuksiin suhteessa alkuperäiseen hilatietoon. Mallin ja hilatiedon erot olivat päätehakkuukuvioilla pieniä. Runkoluvun osalta malli tuotti vähemmän runkoja kuin hilatiedon keskitunnuksissa, mikä voi johtua siitä, että mallin puiden minimiläpimitta alkoi neljästä senttimetristä. Toisaalta referenssikoealoilta ei mitata läpimitaltaan alle viiden senttimetrin puita (Kaukokartoitusperusteisen metsien inventoinnin... 2014), joten neljän senttimetrin läpimittaluokan pitäisi aina olla tyhjä. Hakkuukonetietoon verrattaessa runkoluvun kutistuminen toi mallin keskitunnukset lähemmäksi havaittua. Käytettäessä hakkuukonetietoa havaittuna tietona, voidaan olettaa, etteivät hakkuukoneen mittalaitteen läpi ole kulkeneet alle kuitupuumitan olevat puut. Tämä tarkoittaa, että rinnankorkeusläpimitta on minimissään ollut noin 8-9 senttimetriä. Puustoa on voitu ennen hakkuuta myös ennakkoraivata, jolloin pieniläpimittaiset puut ovat kaadettu ennen hakkuuta

ja jätetty maastoon. Vastaavasti jätetyt säästöpuuryhmät ovat vaikuttaneet suuriläpimittaisten puiden aliarvioon tuotantotiedoissa.

Peuhkurisen ym. (2011) tutkimuksessa puuston kokonaistilavuuden harha oli tätä tutkimusta suurempi, -1,2 % (tässä tutkimuksessa -0,3 %), mutta keskivirhe merkittävästi pienempi, 13,5 % (tässä 17,9 %). Kaikki muut puustotunnukset olivat Peuhkurisen ym. (2011) tutkimuksessa harhattomampia kuin tässä, ja suurin virhe oli runkoluvussa 32,5 % (tässä tutkimuksessa 34,4 %).

Puulajikohtaisesti estimointitulos osoitti männyn yli- ja kuusen aliarviota. Lehtipuiden estimaatti oli sen sijaan harhaton, vaikka sekin sisälsi virhettä. Hilatieto yliarvioi männyn puulajiosuutta 43,4 %, mikä voi johtua alun perin siitä, etteivät laskennassa käytetyt referenssikoealat ole edustaneet tutkimukseen valikoituneita hilasoluja riittävän hyvin. Motti-malli yliarvioi männyn osuutta vieläkin enemmän, 79,9 % suhteessa hakkuukoneaineistoon, mikä näkyy lopullisten tulosten virheissä. Yliarvio oli suurinta niillä kuvioilla, joissa puuston hehtaarikohtainen runkoluku oli suurin. Mallin toimintaa tutkittiin vertailemalla männyn keskitunnuksia eri kuvioilla. Tunnukset G, N, DG ja HG olivat lineaarisesti hyvin riippuvaisia toisistaan, mutta tilavuuden osalta mallin ja hilatiedon keskitunnusten välillä oli selkeä ero. Männyn tilavuuden keskitunnukset olivat keskimäärin 1,23 kertaa suurempia kuin vastaava hilatiedon arvo (Kuva 26). Erikoisinta tässä tapauksessa oli se, ettei mikään muu edellä mainittu puustotunnus viitannut tähän yliarvioon, eikä kuvaajissa näkynyt vastaavaa poikkeamaa. Muilla puulajeillakaan ei vastaavaa yliarviota löytynyt. Kyseessä voi siten olla laskennallinen virhe.



Kuva 26. Männyn keskitilavuuden muutos mallintamisen jälkeen. Muiden männyn puustotunnusten keskiarvo oli lähellä yhtä. Muilla puulajeilla ei esiintynyt vastaavaa poikkeamaa.

5.2 Runkolukusarjat

Runkolukusarjoja voidaan arvioida erilaisin tilastollisin menetelmin ja analyysin. Kun läpimittajakaumia arvioidaan jokin tietyn yhteensopivuustestin avulla, on tärkeää huomioida, mitä testi mittaa. Esimerkiksi Kolmogorov-Smirnov -testin tavoitteena on löytää kumulatiivisen jakauman suurin ero yhden luokan sisällä. Täten testi korostaa hyvin paljon jakauman muotoa, ja yhden luokan merkittävä poikkeus antaa testille heikon arvon riippumatta kyseisen luokan merkittävyydestä. Tilastotieteessä vertaillaan usein empiirisiä jakaumia johonkin tunnettuun jakaumaan, esimerkiksi normaali- tai Weibull-jakaumaan. Testit toimivat hyvin, jos empiirisellä jakaumalla on jokin tarkoitus noudattaa esimerkiksi normaalijakaumaa. Esimerkiksi regressiomallin tehokkuutta voidaan arvioida sillä, kuinka residuaalit asettuvat normaalijakauman tavoin nollan ympärille. Läpimittajakaumissa ei voida tehdä samaa oletusta, sillä suuremmilla läpimittaluokilla on suurempi merkitys puuston kokonaistilavuudessa kuin pienillä.

Tarkastellessa runkolukusarjoja huomattiin, että luokkien väliset suhteelliset erot pienenevät runkoluvun kasvaessa, eli läpimittajakaumat piirtyvät tasaisemmin. K-S testin D-arvo muodostuu kumulatiivisella käyrällä suurimmaksi luokkien määräeroksi, mutta otoksen lukumäärä huomioidaan kriittisen raja-arvon laskennassa (Sokal ja Rohlf 2012).

Suuripuustoisilla kuvioilla myös raja-arvo oli matalampi, mistä johtuu suuripinta-alaisten kuvioiden heikko tilastollinen merkitsevyys, vaikka käyrät näennäisesti olivat lähellä toisiaan. Yksi vaihtoehto olisi ollut laskea raja-arvot puuston hehtaariohjaista lukumäärästä, jolloin pinta-alavaikutus olisi jäänyt huomioimatta. Tulosten tarkastelussa ei kuitenkaan haluttu käyttää hehtaariohjaista arvoja, sillä sen myötä katoaa tietoa kuvion aidosta yhteensopivuudesta. Kuvion pinta-alan kasvun voidaan sanoa lisäävän haluttujen runkojen sekä pituus- ja läpimittasuhteiden löytymisen todennäköisyyttä, mutta se ei välttämättä takaa, että niitä syntyy haluttua määrää. Peuhkurinen ym. (2011) esittivät, ettei virheindeksin laskeminen pienille kuvioille ole tarkoituksenmukaista kuviokoon ollessa pieni. Puuston vähäisen määrän vuoksi on vaikea estimoida kunnollisia runkolukusarjoja, sillä läpimittaluokkien välillä on paljon vaihtelua. Tässä tutkimuksessa jätettiin alle puolen hehtaarin päätehakkuukuviot tutkimuksen ulkopuolelle. Silti etenkin K-S testin tulokset osoittivat, että moni kuvio ei läpäissyt testiä 95 % luottamusvälillä. Tutkimuskuvioiden minimipinta-alaa olisi siten voitu vieläkin kasvattaa, mutta silloin kuviot eivät enää kuvaaisi metsikkökuvioita totuudenmukaisesti; valtaosa kuvioista oli kooltaan alle kaksi hehtaaria. Toisin sanoen metsiköt olivat pienipiirteisiä tällä tutkimusalueella.

Runkolukusarjan sovittaminen yhdellä testillä on haasteellista, mutta tulosten perusteella suurin korrelaatio oli katkonnan lopullisen tarkkuuden ja virheindeksin välillä. Vaikkei virheindeksimenetelmä ole tieteellisesti todistettu, sen hyödynnettävyys on sitä vastoin suuri runkolukusarjojen tutkimisessa.

5.3 Katkonnan tulokset

Katkonnan tulosten tarkkuus noudattaa hyvin aiempien tutkimusten tuloksia. Aiemmista tutkimuksista Siipilehdon ym. (2016) ja Peuhkurisen ym. (2011) oli sovellettu parametrin palautusmenetelmää aluepohjaisella menetelmällä kerättyyn laserkeilaustietoon. Siipilehdon ym. tutkimuksessa katkonnan kokonaistilavuuden harha oli $46,13 \text{ m}^3/\text{ha}$ (13,6 %) (tässä tutkimuksessa $3,8 \text{ m}^3/\text{ha}$ ja 1,5 %) ja keskivirhe $111,3 \text{ m}^3/\text{ha}$ (32,7 %) (tässä tutkimuksessa $42,4 \text{ m}^3/\text{ha}$ ja 16,3 %).

Siipilehdon ym. (2016) tutkimuksessa keskineliövirheen neliöjuuri (RMSE) oli 42,2 % männylle, 106,2 % kuuselle sekä 111,2 % koivulle. Holopaisen ym. (2010) tutkimuksessa kuusen RMSE oli toisaalta vain 33,6 %, mutta pulssitiheys oli tässä aineistossa suurempi, 1,8 pulssia/ m^2 . Sahapuiden tilavuuden RMSE- % sekä harha- % olivat Peuhkurisen ym.

(2008) tutkimuksessa männylle 66,0 % ja -29,6 % sekä kuuselle 31,8 % ja 11,1 %. Holopaisen ym. (2010) tutkimuksessa todettiin tukkipuun ennusteen paranevan suuremmilla läpimittaluokilla, mutta kuitupuun sitä vastoin heikentyvän. Harha oli parhaimmillaan -6,8 % – -12,2 %. Peuhkurisen ym. (2008) tutkimuksessa puustotiedon kokonaistilavuuden laserkeilauksen ja ilmakuvien yhdistelmän RMSE oli 7,0 % sekä harha -0,6 %.

Tutkimuksessa männyn katkonnan suuri prosentuaalinen yliarvio johtui sen vähäisestä määrästä muuhun puustoon nähden. Tuloksissa on havupuulajien sekoittuminen selkeästi havaittavissa tukkipuutavaralajeissa, missä kuusen harhan yliarvio ja männyn aliarvio vastasivat pitkälti toisiaan. Männyn yliarviota lisää vielä se huomio, joka tehtiin keskitunnuksissa; 1,2-kertainen määrän kasvu verrattuna hilatiedon tilavuustunnuksiin vaikuttaa myös katkonnan tuloksissa. Lehtipuiden osalta pieni harha osoittaa, että menetelmä tuottaisi paremman tarkkuuden lehtipuiden estimointiin. Yleisenä johtopäätöksenä voidaan todeta, että puulajien ennustamistarkkuuden heikkous vaikuttaa suoraan puulajikohtaisiin ennustamistarkkuuksiin, mistä johtuen kokonaistilavuuden virhe on tutkimuksissa usein paljon pienempi kuin puulajikohtaiset virheet (Packalen ja Maltamo 2008).

Tutkimusalueena Uudenmaan alue oli hyvin haastava; pienipiirteisiä kuvioita ja epäselviä puulajisuhteita, ja kaukokartoitustiedontuottajan mukaan keilausajankohta ei ollut parhain mahdollinen. Menetelmää tulisi testata laajemmin ympäri eteläistä ja keskeistä Suomea paremman käytettävyyden selvittämiseksi.

6 LOPPUSANAT

Tämän tutkimuksen tarkoituksena oli selvittää Metsäkeskuksen hilatiedon sekä Luonnonvarakeskuksen Motti-mallin käytettävyys runkolukusarjojen sekä puujoukkojen muodostamisessa. Työn aineiston muodostivat 150 päätehakkuukuviota ja tutkimuksen referenssitietona käytettiin kuvioilta kerättyä hakkuukonetietoa. Katkonnan simulointien myötä voitiin tarkastella menetelmän tavaralajeihin asti porautuvaa tarkkuutta.

Runkolukusarjojen tutkiminen antaa hyvän käsityksen siitä, miten puujoukon simulointi on onnistunut. Tässä tutkimuksessa vertailtiin keskenään erilaisia menetelmiä arvioida runkolukusarjojen yhteensopivuutta. Tuloksena selvisi, että ilman pohjapinta-alapainotusta runkolukusarjojen analysointi antaa melko epäkurantteja tuloksia. Vaikka

virheindeksin käyttö ei ole tieteellisesti validoitu menetelmä, sen käytettävyys korrelaatioita etsiessä on ylivoimainen. Kolmogorov-Smirnov -menetelmän käytettävyyttä parantaisi toisaalta runkoluvun jakaminen pinta-alalla sekä arvojen painottaminen pohjapinta-alalla. Uusia menetelmiä runkolukusarjojen vertailuun tarvittaisiin. Tutkimuksessa testattiin runkolukusarjojen arvottamista visuaalisesti. Tapauksessa haluttiin löytää yhteys läpimittafrekvenssihuippujen sekä estimoinnin tarkkuuden välillä. Ne jakaumat, joissa huippuja oli useampi, tarkkuus vaihteli tyypillisesti enemmän, kuin yhden tai kahden huipun sarjoissa.

Katkongan tuloksien virheet sekä harhat vastasivat kokonaistilavuuden osalta paljon aiempia tutkimuksia. Männyn harhat sekä virheet johtuivat hilatiedon ja Motti-mallin puulajin yliarviosta, joka puolestaan vaikutti kuusen aliarvioon. Testeissä ei löydetty syytä, miksi malli yliarvioi männyn määrää. Toisaalta männyn osuus oli tutkimusaineistossa pieni, joten estimointivirheelle tulisi antaa myös sen mukainen painoarvo.

Runkolukusarjainen metsävaratieto ehdottomasti parantaisi puunhankinnan estimoinnin tarkkuutta, ja tulevaisuudessa vaikutus tulee näkymään puunhankinnan strategisessa päätöksenteossa. Puuston laadun estimointi jäi tämän tutkimuksen ulkopuolelle. Vaikka aiemmissa tutkimuksissa laadun estimointiin on käytetty valtakunnallisia tukkivähennysmalleja (Mehtätalo 2002), ei niillä voida selittää pienipiirteisiä eroja metsikkökuvioiden välillä (Malinen ym. 2011). Esimerkiksi männyn kuivaoksaraja on merkittävä laadun estimaattori ja metsänhoidollisilla tekijöillä on paljon vaikutusta kuivaoksarajan korkeuteen. Laadun huomioonotto vaatii erillisiä mittauksia mm. kuivaoksarajan suhteen. Hakkuukoneaineiston käyttö ”laatuselittäjänä” ei myöskään olisi onnistunut, sillä laatukatkonnassa taustalla vaikuttaa myös tavoitejakauma; jokainen A- ja B-tukin katkaisu ei johdu laatumuutoksesta rungossa. Kuusen tyvilahon estimointi on haasteellista vähäisen tutkimuksen vuoksi. Hakkuukone ei myöskään tallenna leikkoja, eli tyvestä poistettuja lahoa sisältäviä rungon viipaleita järjestelmään.

Tutkimuksessa keskitunnusten avulla tehdyn analyysin perusteella merkittävien virhetekijä oli hilatieto. Kokojakaumamalli noudatti hilatiedon keskitunnuksia kohtuullisissa rajoissa. Hilatiedon runkolukusarjojen ennustamiseen on kokojakaumamallilla hyvät edellytykset, mutta tutkimuksessa käytetyn yhdistelmän tarkkuutta tulisi tutkia laajemmassa testiympäristössä.

KIRJALLISUUS

Alliez, P., Delage, C., Karavelas, M. I., Pion, S., Teillaud, M., ja Yvinec, M. **2009**. Delaunay Tessellations and Voronoi Diagrams in CGAL. *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*. Springer.

Armstrong, J. S. 1978. *Long-range Forecasting: From Crystal Ball to Computer*. Wiley. 612 s.

Gove J.H., Patil G.P. (1998). Modeling the basal area-size distribution of forest stand: a compatible approach. *Forest Science* 44(2): 285–297.

Haara, A. & Korhonen, K.T. 2004. Kuvioittaisen arvioinnin luotettavuus. *Metsätieteen aikakauskirja* 4/2004: 489–508. (In Finnish).

Holopainen M., Vastaranta M., Rasinmäki J., Kalliovirta J., Mäkinen A., Haapanen R., Melkas T., Yu X., Hyypä J. 2010. Uncertainty in timber assortment estimates predicted from forest inventory data. *European Journal of Forest Research* 129: 1131–1142.

Hudak A.T., Crookston N.L., Evans J.S., Hall D.E., Falkowski M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment* 112(5): 2232–2245.
<http://dx.doi.org/10.1016/j.rse.2007.10.009>.

Hyink D.M., Moser J.W. (1983). A generalized framework for projecting forest yield and stand structure using diameter distributions. *Forest Science* 29: 85–95.

Hämäläinen, J., Räsänen, T., Ritala, R., Häme, T., Tergujeff, R. 2017 Seuraavan sukupolven palvelualustan kehittäminen metsätiedon jakeluun. *Metsätehon raportti* 241. ISSN 1796-2374 (Verkkojulkaisu).

Junttila, V., Maltamo, M., Kauranne, T. 2008. Sparse Bayesian Estimation of Forest Stand Characteristics from Airborne Laser Scanning. *Forest Science* 54(5): 543 - 552.

Kaukokartoitusperusteisen metsien inventoinnin koalojen maastotyöohje. 2014. *Maastotyöohje*. Suomen Metsäkeskus. 29 s.

Kivinen, V.-P. 2004. A Genetic Algorithm Approach to Tree Bucking Optimization. *Forest Science* 50(5): 696-710.

- , Uusitalo, J. 2002. Applying fuzzy logic to tree bucking control. *Forest Science* 48(4): 673–684.

- Korhonen L., Peuhkurinen J., Malinen J., Maltamo M., Suvanto A., Packalén P., Kangas J. 2008. The use of airborne laser scanning to estimate sawlog volumes. *Forestry* 81: 499–510.
- Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce and birch. *Communicationes Instituti Forestalis Fenniae* 108. 74 s.
- Lourakis, M. I. A. 2005. A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar. Foundation of Research and Technology, Hellas. Saatavissa: <http://cpansearch.perl.org/src/JLAPEYRE/PDL-Fit-Levmar-0.0091/doc/levmar.pdf>. [Viitattu 16.7.2017]
- Malinen, J., Wall, T., Kilpeläinen, H., Verkasalo, E. 2011. Leimikon arvonmuodostus vaihtoehtoisissa loppukäyttökohteissa. *Metlan työraportteja* 206. ISBN 978-951-40-2315-6.
- Maltamo, M., Kangas, A. 1998. Methods based on k-nearest neighbor regression in the prediction of basal area diameter distribution. *Canadian Journal of Forest Research*. 28(8): 1107-1115, 10.1139/x98-085
- Maltamo, M., Haara, A., Hirvelä, H., Kangas, A., Lempinen, R., Malinen, J., Nalli, A., Nuutinen, T. & Siipilehto, J. 2002. Lämpimittajakaumamalleihin perustuvat vaihtoehdot kuvauspuiden muodostamiseen puuston keskitunnustietojen avulla. *Metsätieteen aikakauskirja* 3/2002: 407–423.
- Mehtätalo, L. 2002. Valtakunnalliset puukohtaiset tukkivähennysmallit männyille, kuuselle, koivuille ja haavalle. *Metsätieteen aikakauskirja* 4/2002: 575–591.
- Melkas, T. 2017. Henkilökohtainen tiedonanto.
- Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* 80(1): 88–99. doi:10.1016/S0034-4257(01)00290-5.
- Naesset, E. 2004. Accuracy of Forest Inventory Using Airborne Laser scanning: Evaluating the First Nordic Full-scale Operational Project. *Scandinavian Journal of Forest Research* 19: 554 – 557.
- Nissinen, J. 2002. Improving compatibility between prediction of basal area diameter distributions and assessments of young stands. *Metsäsuunnittelun ja -ekonomian pro gradu*. University of Joensuu. 48 p.
- Packalén, P., Maltamo M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Canadian Journal of Forest Research* 38: 1750–1760.

- Packalén, P., Maltamo, M. 2007. The k-MSN method in the prediction of species specific stand attributes using airborne laser scanning and aerial photographs. *Remote Sensing of Environment* 109: 328–341
- Peuhkurinen, J., Maltamo M., Malinen J. 2008. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: A distribution-based approach. *Silva Fennica* 42: 625–641.
- Peuhkurinen, J., Mehtätalo L., Maltamo M. 2011. Comparing individual tree detection and the area-based statistical approach for retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. *Canadian Journal of Forest Research* 41(3): 583–598. <http://dx.doi.org/10.1139/X10-223>.
- Piira, T., Kilpeläinen, H., Malinen, J., Wall, T. & Verkasalo, E. 2007. Leimikon puutavaralajikertymän ja myyntiarvon vaihtelu erilaisilla katkontaohjeilla. *Metsätieteen aikakauskirja* 1/2007: 19–37.
- Poikela, A. 2017. Henkilökohtainen tiedonanto.
- Poso, S. 2001. Metsänarvioimistiede osana suomalaista tietoyhteiskuntaa. *Metsätieteen aikakauskirja* 2/2001: 290-295.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Saatavissa: <http://www.R-project.org/>. [Viitattu 17.9.2017]
- Riekkilä, K. 2017. Henkilökohtainen tiedonanto.
- Siipilehto, J. 2011a. Local prediction of stand structure using linear prediction theory in Scots pine-dominated stands in Finland. *Silva Fennica* 45(4): 669–692.
- . 2011b. Methods and applications for improving parameter prediction models for stand structures in Finland. *Dissertationes Forestales* 124.
- , Mehtätalo L. 2013. Parameter recovery vs. parameter prediction for the Weibull distribution validated for Scots pine stands in Finland. *Silva Fennica* vol. 47 no. 4 article id 1057. 22 p.
- , Kangas, A. 2015. Näslundin pituuskäyrä ja siihen perustuvia malleja läpimitan ja pituuden välisestä riippuvuudesta suomalaisissa talousmetsissä. *Metsätieteen aikakauskirja* 4/2015: 215–236.
- , Lindeman H., Vastaranta M., Yu X., Uusitalo J. 2016. Reliability of the predicted stand structure for clear-cut stands using optional methods: airborne laser scanning-based methods, smartphone-based forest inventory application Trestima and pre-harvest

measurement tool EMO. *Silva Fennica* vol. 50 no. 3 article id 1568. 24 p.
<http://dx.doi.org/10.14214/sf.1568>.

Sokal R.R., Rohlf F.J. (1981). *Biometry. The principles and practice of statistics in biological research*. W.H. Freeman and Company, San Francisco. 843 p.

StanFord. 2016. Standard for forestry data and communication. SkogForsk. [www-sivusto]. Saatavissa: <http://www.skogforsk.se/english/projects/stanford/>. [Viitattu 17.9.2017]

Suomen metsäkeskuksen metsävaratiedon laatuseloste. 2016. Suomen metsäkeskus. [Verkkodokumentti]. Saatavissa: www.metsakeskus.fi/sites/default/files/metsavaratiedon_laatuseloste.pdf. [Viitattu 17.9.2017]

Suvanto, A., Maltamo, M., Packalén, P. & Kangas, J. 2005. Kuviokohtaisten puustotunnusten ennustaminen laserkeilauksella. *Metsätieteen aikakauskirja* 4/2005: 413–428

Terminen kasvukausi. 2017. Vuodenaikojen tilastot. Ilmatieteen laitos. [www-sivu]. Saatavissa: <http://ilmatieteenlaitos.fi/terminen-kasvukausi> [Viitattu 16.7.2017]

Tham Å. 1988. Structure of Mixed *Picea abies* (L.) Karst. and *Betula pendula* Roth and *Betula pubescens* Ehrh. stands in South and middle Sweden. *Scandinavian Journal of Forest Research* 3(1–4): 355–369. <http://dx.doi.org/10.1080/02827588809382523>.

Tofallis, C. 2015. A Better Measure of Relative Prediction Accuracy for Model Selection and Model Estimation. *Journal of the Operational Research Society*. 66(8) s. 1352-1362

Uusitalo, J. 1997. Pre-harvest measurement of pine stands for sawing production planning. University of Helsinki, Department of Forest Resource Management, Publications 9. 96 p.

- , Kokko, S. & Kivinen, V.-P. 2004. The effect of two bucking methods on Scots pine lumber quality. *Silva Fennica* 38(3): 291–303.

Reynolds, M.R., Jr., Burk, T.E. & Huang, W.-C. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *Forest Science* 34(2): 373–399.

Quantum GIS Development Team. 2017. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. [tietokoneohjelma] Saatavissa: <http://qgis.osgeo.org>